PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/72 // 14/585 — A1

(11) International Publication Number: **WO 98/21242**

(43) International Publication Date: 22 May 1998 (22.05.98)

(21) International Application Number: PCT/US97/21330

(22) International Filing Date: 14 November 1997 (14.11.97)

(30) Priority Data:
60/030,934 15 November 1996 (15.11.96) US

(71) Applicant: HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).

(72) Inventors: SOPPET, Daniel, R.; 15050 Stillfield Place, Centreville, VA 22020 (US). LI, Yi; 16125 Howard Landing Drive, Gaithersburg, MD 20878 (US). RUBEN, Steven, M.; 18528 Heritage Hills Drive, Olney, MD 20832 (US).

(74) Agents: STEFFE, Eric, K. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005–3934 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.
With an indication in relation to a deposited microorganism furnished under Rule 13bis separately from the description.
Date of receipt by the International Bureau:
23 December 1997 (23.12.1997)

(54) Title: CALCITONIN RECEPTOR

(57) Abstract

The present invention relates to a novel calcitonin receptor protein which is a member of the G–protein coupled receptor superfamily. In particular, isolated nucleic acid molecules are provided encoding the human calcitonin receptor protein. Calcitonin receptor polypeptides are also provided as are vectors, host cells and recombinant methods for producing the same. The invention further relates to screening methods for identifying agonists and antagonists of calcitonin receptor activity. Also disclosed are diagnostic assays for detecting diseases associated with altered calcitonin receptor expression.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Calcitonin Receptor

Background of the Invention

Field of the Invention

The present invention relates to a novel human G-protein coupled receptor. More specifically, isolated nucleic acid molecules are provided encoding a human calcitonin receptor. Calcitonin receptor polypeptides are also provided, as are vectors, host cells and recombinant methods for producing the same. The invention further relates to screening methods for identifying agonists and antagonists of calcitonin receptor activity. Also disclosed are diagnostic assays for detecting diseases associated with altered calcitonin receptor expression.

Related Art

It is well established that many medically significant biological processes are mediated by proteins participating in signal transduction pathways that involve G-proteins and/or second messengers, e.g., cAMP (Lefkowitz, *Nature, 351*:353-354 (1991)). Herein these proteins are referred to as proteins participating in pathways with G-proteins or PPG proteins. Some examples of these proteins include the GPC receptors, such as those for adrenergic agents and dopamine (Kobilka, B.K., *et al., PNAS, 84*:46-50 (1987); Kobilka, B.K., *et al., Science, 238*:650-656 (1987); Bunzow, J.R., *et al., Nature, 336*:783-787 (1988)), G-proteins themselves, effector proteins, e.g., phospholipase C, adenyl cyclase, and phosphodiesterase, and actuator proteins, e.g., protein kinase A and protein kinase C (Simon, M.I., *et al., Science, 252*:802-808 (1991)).

For example, in one form of signal transduction, the effect of hormone binding is activation of an enzyme, adenylate cyclase, inside the cell. Enzyme activation by hormones is dependent on the presence of the nucleotide GTP, and GTP also influences hormone binding. A G-protein connects the hormone receptors to adenylate cyclase. G-protein was shown to exchange GTP for bound GDP when activated by hormone receptors. The GTP-carrying form then binds to an activated adenylate cyclase. Hydrolysis of GTP to GDP, catalyzed by the G-protein itself, returns the G-protein to its basal, inactive form. Thus, the G-protein serves a dual role, as an intermediate that relays the signal from receptor to effector, and as a clock that controls the duration of the signal.

The membrane protein gene superfamily of G-protein coupled receptors has been characterized as having seven putative transmembrane domains. The domains are believed to represent transmembrane α-helices connected by extracellular or cytoplasmic loops. G-protein coupled receptors include a wide range of biologically active receptors, such as hormone, viral, growth factor and neuroreceptors.

G-protein coupled receptors have been characterized as including these seven conserved hydrophobic stretches of about 20 to 30 amino acids, connecting at least eight divergent hydrophilic loops. The G-protein family of coupled receptors includes dopamine receptors which bind neuroleptic drugs used for treating psychotic and neurological disorders. Other examples of members of this family include calcitonin, adrenergic, endothelin, cAMP, adenosine, muscarinic, acetylcholine, serotonin, histamine, thrombin, kinin, follicle stimulating hormone, opsins, endothelial differentiation gene-1 receptor, rhodopsins, odorant, cytomegalovirus receptors, etc.

Most G-protein coupled receptors have single conserved cysteine residues in each of the first two extracellular loops which form disulfide bonds that are believed to stabilize functional protein structure. The 7 transmembrane regions are designated as TM1, TM2, TM3, TM4, TM5, TM6, and TM7. TM3 has been implicated in signal transduction.

Phosphorylation and lipidation (palmitylation or farnesylation) of cysteine residues can influence signal transduction of some G-protein coupled receptors. Most G-protein coupled receptors contain potential phosphorylation sites within the third cytoplasmic loop and/or the carboxy terminus. For several G-protein coupled receptors, such as the β-adrenoreceptor, phosphorylation by protein kinase A and/or specific receptor kinases mediates receptor desensitization.

For some receptors, the ligand binding sites of G-protein coupled receptors are believed to comprise a hydrophilic socket, formed by several G-protein coupled receptors transmembrane domains, that is surrounded by hydrophobic residues of the G-protein coupled receptors. The hydrophilic side of each G-protein coupled receptor transmembrane helix is postulated to face inward and form the polar ligand binding site. TM3 has been implicated in several G-protein coupled receptors as having a ligand binding site, including the TM3 aspartate residue. Additionally, TM5 serines, a TM6 asparagine and TM6 or TM7 phenylalanines or tyrosines are also implicated in ligand binding.

G-protein coupled receptors can be intracellularly coupled by heterotrimeric G-proteins to various intracellular enzymes, ion channels and transporters (see, Johnson *et al.*, *Endoc., Rev., 10*:317-331 (1989)). Different G-protein α-subunits preferentially stimulate particular effectors to modulate various biological functions in a cell. Phosphorylation of cytoplasmic residues of G-protein coupled receptors have been identified as an important mechanism for the regulation of G-protein coupling of some G-protein coupled receptors. G-protein coupled receptors are found in numerous sites within a mammalian host.

Over the past 15 years, nearly 150 therapeutic agents targeting 7 transmembrane (7 TM) receptors have been successfully introduced onto the market. This indicates that these receptors have an established, proven history as therapeutic targets. Clearly there is a need for identification and characterization of further receptors which can play a role in preventing,

ameliorating or correcting dysfunctions or diseases, including, but not limited to, pain, cancers, anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, myocardial infarction, ulcers, asthma, allergies, benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation, dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others.

Summary of the Invention

The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding the calcitonin receptor having the amino acid sequence shown in Figures 1 A and B (SEQ ID NO:2) or the amino acid sequence encoded by the cDNA clone deposited in a bacterial host as ATCC Deposit Number 97730 on September 23, 1996.

The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of calcitonin receptor polypeptides or peptides by recombinant techniques.

The invention further provides an isolated calcitonin polypeptide having an amino acid sequence encoded by a polynucleotide described herein. The present invention also provides a screening method for identifying compounds capable of enhancing or inhibiting (i.e., agonists or antagonists) a cellular response induced by the calcitonin receptor.,

The invention further provides methods for isolating antibodies that bind specifically to a calcitonin receptor polypeptide having an amino acid sequence as described herein. Such antibodies are useful diagnostically or therapeutically as describe below.

G-protein coupled receptors, such as the calcitonin receptor of the present invention, play a role a number of dysfunctions or diseases. Thus, agonists or antagonists of the calcitonin receptor of the present invention could be used therapeutically to prevent, ameliorate or correct dysfunctions or diseases associated with the calcitonin receptor of the present invention. Another aspect of the invention is related to methods for treating an individual in need of an increased or decreased level of calcitonin receptor activity in the body.

Brief Description of the Figures

Figures 1 A and B show the nucleotide (SEQ ID NO:1) and deduced amino acid (SEQ ID NO:2) sequences of human calcitonin receptor. The protein has a deduced molecular weight of about 61287 Da.

Figure 2 shows the predicted location of each of the seven transmembrane domains of the calcitonin receptor.

Detailed Description

The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding a calcitonin receptor polypeptide having the amino acid sequence shown in Figures 1 A and B (SEQ ID NO:2), which was determined by sequencing a cloned cDNA. The calcitonin receptor protein of the present invention shares sequence homology with rat calcitonin receptor A, rat calcitonin receptor B, rat growth hormone releasing hormone receptor GHRH, and hormone receptor EMR1 (SEQ ID NOs:3-6). The nucleotide sequence shown in Figures 1 A and B (SEQ ID NO:1) was obtained by sequencing the HCEPR64 clone, which was deposited on September 23, 1996 at the American

Type Culture Collection, 12301 Park Lawn Drive, Rockville, Maryland 20852, and given accession number 97730. The deposited clone is contained in the pBluescript SK(-) plasmid (Stratagene, La Jolla, CA).

Nucleic Acid Molecules

Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer (such as the Model 373 from Applied Biosystems, Inc.), and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

Using the information provided herein, such as the nucleotide sequence in Figures 1 A and B, a nucleic acid molecule of the present invention encoding a calcitonin receptor polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Illustrative of the invention, the nucleic acid molecule described in Figures 1 A and B (SEQ ID NO:1) was discovered in a cDNA library derived

from cerebellum using the expressed sequence tag (EST) analysis (Adams, *et al. Science 252*:1651-1656 (1991); Adams *et al. Nature 355*:632-634(1992); Adams *et al. Nature 377* Supp, 3-174 (1995)). The gene was also identified in cDNA libraries from the following tissues cerebellum, early brain, fetal brain and uterine cancer. The determined nucleotide sequence of the calcitonin receptor cDNA of Figures 1 A and B (SEQ ID NO:1) contains an open reading frame encoding a protein of 568 amino acid residues, with an initiation codon at positions 652-654 of the nucleotide sequence in Figures 1 A and B (SEQ ID NO:1), and a deduced molecular weight of about 61287 Da. The calcitonin receptor protein shown in Figures 1 A and B (SEQ ID NO:2) is about 21% identical to rat calcitonin receptor A, is about 21% identical to rat calcitonin receptor B, is about 25% identical to rat growth hormone releasing hormone receptor GHRH, and is about 31% identical to hormone receptor EMR1.

As one of ordinary skill would appreciate, due to the possibilities of sequencing errors discussed above, the actual calcitonin receptor polypeptide encoded by the deposited cDNA comprises about 568 amino acids, but may be anywhere in the range of 555-575 amino acids.

As indicated, nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA

molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

Isolated nucleic acid molecules of the present invention include DNA molecules comprising an open reading frame (ORF) shown in Figures 1 A and B (SEQ ID NO:1); DNA molecules comprising the coding sequence for the calcitonin receptor protein shown in Figures 1 A and B (568 amino acids) (SEQ ID NO:2); and DNA molecules which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode the calcitonin receptor protein. Of course, the genetic code is well known in the art. Thus, it would be routine for one skilled in the art to generate the degenerate variants described above.

In addition, the invention provides nucleic acid molecules having nucleotide sequences related to extensive portions of SEQ ID NO:1 which have been determined from the following related cDNA clones: HCEPR64R (SEQ ID NO:13), and HCETE57R (SEQ ID NO:14).

The sequence of a public ESTs also having nucleotide sequences related to extensive portions of SEQ ID NO:1, are the following GenBank Accession Nos: R60111 (SEQ ID NO:15), R60049 (SEQ ID NO:16), AA326354 (SEQ ID NO:17), W52864 (SEQ ID NO:18), AA037392 (SEQ ID NO:19), T06233 (SEQ ID NO:20), AA477598 (SEQ ID NO:21), F13341 (SEQ ID NO:22), AA477470 (SEQ ID NO:23), T77092 (SEQ ID NO:24), H55526 (SEQ ID NO:25), and AA349458 (SEQ ID NO:26).

In another aspect, the invention provides isolated nucleic acid molecules encoding the calcitonin receptor polypeptide having an amino acid sequence encoded by the cDNA clone contained in the plasmid deposited as ATCC Deposit No. 97730 on September 23, 1996. The invention further provides an isolated nucleic acid molecule having the nucleotide sequence shown in Figures 1 A and B (SEQ ID NO:1) or the nucleotide sequence of the calcitonin receptor cDNA contained in the above-described deposited clone, or a nucleic acid molecule having a sequence complementary to one of the above sequences. Such isolated

molecules, particularly DNA molecules, are useful as probes for gene mapping, by *in situ* hybridization with chromosomes, and for detecting expression of the calcitonin receptor gene in human tissue, for instance, by Northern blot analysis.

The present invention is further directed to fragments of the isolated nucleic acid molecules described herein. By a fragment of an isolated nucleic acid molecule having the nucleotide sequence of the deposited cDNA or the nucleotide sequence shown in Figures 1 A and B (SEQ ID NO:1) is intended fragments at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length which are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000, 1025, 1050, 1075, 1100, 1125, 1150,1175, 1200, 1225, 1250, 1275, 1300, 1325, 1350, 1375, 1400, 1425, 1450, 1475 or 1500 nt in length are also useful according to the present invention as are fragments corresponding to most, if not all, of the nucleotide sequence of the deposited cDNA or as shown in Figures 1 A and B (SEQ ID NO:1). By a fragment at least 20 nt in length, for example, is intended fragments which include 20 or more contiguous bases from the nucleotide sequence of the deposited cDNA or the nucleotide sequence as shown in Figures 1 A and B (SEQ ID NO:1).

Preferred nucleic acid fragments of the present invention include nucleic acid molecules encoding epitope-bearing portions of the calcitonin receptor protein that were predicted from the results of an "Antigenic Index-Jameson-Wolf" graph. In particular, such nucleic acid fragments of the present invention include nucleic acid molecules encoding: a polypeptide comprising amino acid residues from about 10 to about 20 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 49 to about 60 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 113 to about 123 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 145 to about 154 in

SEQ ID NO:2; a polypeptide comprising amino acid residues from about 189 to about 209 in SEQ ID NO:2; and a polypeptide comprising amino acid residues from about 259 to about 560 in SEQ ID NO:2. The inventors have determined that the above polypeptide fragments are antigenic regions of the calcitonin receptor protein. Methods for determining other such epitope-bearing portions of the calcitonin receptor protein are described in detail below.

In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention described above, for instance, the cDNA clone contained in ATCC Deposit No. 97730. By "stringent hybridization conditions" is intended overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150 mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 g/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

By a polynucleotide which hybridizes to a "portion" of a polynucleotide is intended a polynucleotide (either DNA or RNA) hybridizing to at least about 15 nucleotides (nt), and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably about 30-70 nt of the reference polynucleotide. These are useful as diagnostic probes and primers as discussed above and in more detail below.

By a portion of a polynucleotide of "at least 20 nt in length," for example, is intended 20 or more contiguous nucleotides from the nucleotide sequence of the reference polynucleotide (e.g., the deposited cDNA or the nucleotide sequence as shown in Figures 1 A and B (SEQ ID NO:1)). Of course, a polynucleotide which hybridizes only to a poly A sequence (such as the 3′ terminal poly(A) tract of the calcitonin receptor cDNA shown in Figures 1 A and B (SEQ ID NO:1)), or to a complementary stretch of T (or U) resides, would not be included in a polynucleotide of the invention used to hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid

molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

As indicated, nucleic acid molecules of the present invention which encode a calcitonin receptor polypeptide may include, but are not limited to those encoding the coding sequence for the polypeptide and additional sequences, such as a pre-, or pro- or prepro- protein sequence; the coding sequence of the polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences, including for example, but not limited to introns and non-coding 5′ and 3′ sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing, including splicing and polyadenylation signals, for example - ribosome binding and stability of mRNA; an additional coding sequence which codes for additional amino acids, such as those which provide additional functionalities. Thus, the sequence encoding the polypeptide may be fused to a marker sequence, such as a sequence encoding a peptide which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (Qiagen, Inc.), among others, many of which are commercially available. As described in Gentz *et al., Proc. Natl. Acad. Sci. USA 86*:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. The "HA" tag is another peptide useful for purification which corresponds to an epitope derived from the influenza hemagglutinin protein, which has been described by Wilson *et al., Cell 37*: 767 (1984). As discussed below, other such fusion proteins include the calcitonin receptorfused to Fc at the N- or C-terminus.

The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of the calcitonin receptor protein. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. *Genes II*,

Lewin, B., ed., John Wiley & Sons, New York (1985). Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

Such variants include those produced by nucleotide substitutions, deletions or additions, which may involve one or more nucleotides. The variants may be altered in coding regoins, non-coding regions, or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the calcitonin receptor protein or portions thereof. Also especially preferred in this regard are conservative substitutions.

Further embodiments of the invention include isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical to (a) a nucleotide sequence encoding the calcitonin receptor polypeptide having the amino acid sequence in SEQ ID NO:2; (b) a nucleotide sequence encoding the calcitonin receptor polypeptide having the amino acid sequence in SEQ ID NO:2, but lacking the N-terminal methionine; (c) a nucleotide sequence encoding the calcitonin receptor polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97730; (d) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor extracellular regions selected from the group consisting of: amino acid residues from about 1 to about 18, amino acid residues from about 148 to about 159, amino acid residues from about 188 to about 209, and amino acid residues from about 231 to about 236 in SEQ ID NO:2; (e) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor intracellular regions selected from the group consisting of:amino acid residues from about 50 to about 59, amino acid residues from about 110 to about 125, amino acid residues from about 188 to about 209, and amino acid residues from about 259 to about 560 in SEQ ID NO:2; (f) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 1 (amino acid residues from about 20 to about 49 in SEQ

ID NO:2); (g) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 2 (amino acid residues from about 60 to about 80 in SEQ ID NO:2); (h) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 3 (amino acid residues from about 82 to about 109 in SEQ ID NO:2); (i) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 4 (amino acid residues from about 126 to about 147 in SEQ ID NO:2); (j) a nucleotide sequence encoding the calcitonin receptor transmembrane domain 5 (amino acid residues from about 160 to about 187 in SEQ ID NO:2); (k) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 6 (amino acid residues from about 209 to about 230 in SEQ ID NO:2); (l) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 7 (amino acid residues from about 237 to about 258 in SEQ ID NO:2); and (m) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), (h), (i), (j), (k) or (l), above.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding a calcitonin receptorpolypeptide is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the calcitonin receptorpolypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5′ or 3′ terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually

among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular nucleic acid molecule is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the nucleotide sequence shown in Figures 1 A and B or to the nucleotides sequence of the deposited cDNA clone can be determined conventionally using known computer programs such as the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711. Bestfit uses the local homology algorithm of Smith and Waterman, *Advances in Applied Mathematics 2*: 482-489 (1981), to find the best segment of homology between two sequences. When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed.

The present application is directed to nucleic acid molecules at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figures 1 A and B (SEQ ID NO:1) or to the nucleic acid sequence of the deposited cDNA, irrespective of whether they encode a polypeptide having calcitonin receptor activity. This is because even where a particular nucleic acid molecule does not encode a polypeptide having calcitonin receptor activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe or a polymerase chain reaction (PCR) primer. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having calcitonin receptor activity include, *inter alia,* (1) isolating the calcitonin receptor gene or allelic variants thereof in a cDNA library; (2) *in situ* hybridization (e.g., "FISH") to metaphase chromosomal spreads to provide precise chromosomal location of the calcitonin receptor gene, as described in

Verma *et al., Human Chromosomes: A Manual of Basic Techniques*, Pergamon Press, New York (1988); and Northern Blot analysis for detecting calcitonin receptor mRNA expression in specific tissues.

Preferred, however, are nucleic acid molecules having sequences at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figures 1 A and B (SEQ ID NO:1) or to the nucleic acid sequence of the deposited cDNA which do, in fact, encode a polypeptide having calcitonin receptor protein activity. By "a polypeptide having calcitonin receptor activity" is intended polypeptides exhibiting activity similar, but not necessarily identical, to an activity of the calcitonin receptor protein of the invention, as measured in a particular biological assay.

One can evaluate whether a candidate polypeptide has calcitonin receptor activity by performing simple binding kinetics to measure receptor affinity for a particular ligand (i.e., calcitonin). Binding kinetic analysis experiments are well known to those skilled in the art (Chizzonite *et al., Proc Natl. Acad. Sci. 86*:8029-8033 (1989); Mizel, *et al., J. Immunol. 138*:2906-2912 (1987)). One can monitor the presence of calcitonin receptor activity by testing whether the putative calcitonin receptor polypeptide has a similar binding affinity for a particular ligand as the receptor of the present invention. The receptor of the present invention can serve as a reference for the assay for receptor activity.

Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 90%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid sequence of the deposited cDNA or the nucleic acid sequence shown in Figures 1 A and B (SEQ ID NO:1) will encode a polypeptide "having calcitonin receptor protein activity." In fact, since degenerate variants of these nucleotide sequences all encode the same polypeptide, this will be clear to the skilled artisan even without performing the above described comparison assay. It will be further recognized in the art that, for such nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a

polypeptide having calcitonin receptor protein activity. This is because the skilled artisan is fully aware of amino acid substitutions that are either less likely or not likely to significantly effect protein function (e.g., replacing one aliphatic amino acid with a second aliphatic amino acid).

For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. *et al.*, "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," *Science 247:*1306-1310 (1990), wherein the authors indicate that proteins are surprisingly tolerant of amino acid substitutions.

Vectors and Host Cells

The present invention also relates to vectors which include the isolated DNA molecules of the present invention, host cells which are genetically engineered with the recombinant vectors, and the production of calcitonin receptor polypeptides or fragments thereof by recombinant techniques.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged *in vitro* using an appropriate packaging cell line and then transduced into host cells.

The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli lac, trp* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture and tetracycline or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis *et al.*, *Basic Methods In Molecular Biology* (1986).

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide

moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art. A preferred fusion protein comprises a heterologous region from immunoglobulin that is useful to solubilize proteins. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is thoroughly advantageous for use in therapy and diagnosis and thus results, for example, in improved pharmacokinetic properties (EP-A 0232 262). On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified in the advantageous manner described. This is the case when Fc portion proves to be a hindrance to use in therapy and diagnosis, for example when the fusion protein is to be used as antigen for immunizations. In drug discovery, for example, human proteins, such as, hIL5-receptor has been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. See, D. Bennett *et al.*, *Journal of Molecular Recognition*, Vol. 8:52-58 (1995) and K. Johanson *et al.*, *The Journal of Biological Chemistry*, Vol. 270, No. 16:9459-9471 (1995).

The calcitonin receptor protein can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a

recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

Calcitonin Receptor Polypeptides and Fragments

The invention further provides an isolated calcitonin receptor having the amino acid sequence encoded by the deposited cDNA, or the amino acid sequence in Figures 1 A and B (SEQ ID NO:2), or a peptide or polypeptide comprising a portion of the above polypeptides.

It will be recognized in the art that some amino acid sequences of the calcitonin receptor polypeptide can be varied without significant effect of the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine activity.

Thus, the invention further includes variations of the calcitonin receptor polypeptide which show substantial calcitonin receptor activity or which include regions of calcitonin receptor protein such as the protein portions discussed below. Such mutants include deletions, insertions, inversions, repeats, and type substitutions. As indicated above, guidance concerning which amino acid changes are likely to be phenotypically silent can be found in Bowie, J.U., *et al.*, "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," *Science 247:*1306-1310 (1990).

Thus, the fragment, derivative or analog of the polypeptide of Figures 1 A and B (SEQ ID NO:2), or that encoded by the deposited cDNA, may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes

a substituent group, or (iii) one in which the polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the polypeptide, such as an IgG Fc fusion region peptide or leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

Of particular interest are substitutions of charged amino acids with another charged amino acid and with neutral or negatively charged amino acids. The latter results in proteins with reduced positive charge to improve the characteristics of the calcitonin protein. The prevention of aggregation is highly desirable. Aggregation of proteins not only results in a loss of activity but can also be problematic when preparing pharmaceutical formulations, because they can be immunogenic. (Pinckard *et al., Clin Exp. Immunol. 2*:331-340 (1967); Robbins *et al., Diabetes 36*:838-845 (1987); Cleland *et al. Crit. Rev. Therapeutic Drug Carrier Systems 10*:307-377 (1993)).

As indicated, changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein (see Table 1).

TABLE 1. Conservative Amino Acid Substitutions.

Aromatic	Phenylalanine Tryptophan Tyrosine
Hydrophobic	Leucine Isoleucine Valine
Polar	Glutamine Asparagine
Basic	Arginine Lysine Histidine
Acidic	Aspartic Acid Glutamic Acid
Small	Alanine Serine Threonine Methionine Glycine

Of course, the number of amino acid substitutions a skilled artisan would make depends on many factors, including those described above. Generally speaking, the number of substitutions for any given calcitonin receptor polypeptide will not be more than 50, 40, 30, 25, 20, 15, 10, 5 or 3.

Amino acids in the calcitonin receptor protein of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, *Science 244*:1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as receptor binding or *in vitro*, or *in vitro* proliferative activity. Sites that are critical for ligand-receptor binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith *et al., J. Mol. Biol. 224*:899-904 (1992) and de Vos *et al. Science 255*:306-312 (1992)).

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. By "isolated polypeptide"

is intended a polypeptide removed from its native environment. Thus, a polypeptide produced and/or contained within a recombinant host cell is considered isolated for purposes of the present invention. Also intended as an "isolated polypeptide" are polypeptides that have been purified, partially or substantially, from a recombinant host cell. For example, a recombinantly produced version of the calcitonin receptor polypeptide can be substantially purified by the one-step method described in Smith and Johnson, *Gene 67:*31-40 (1988).

The polypeptides of the present invention include the polypeptide encoded by the deposited cDNA, the polypeptide encoded by the deposited cDNA minus the N-terminal methionine, a polypeptide comprising amino acids about 1 to about 568 in SEQ ID NO:2, a polypeptide comprising amino acids about 2 to about 568 in SEQ ID NO:2, the extracellular regions, the intracellular regions, transmembrane domain 1, transmembrane domain 2, transmembrane domain 3, transmembrane domain 4, transmembrane domain 5, transmembrane domain 6, transmembrane domain 7, as well as polypeptides which are at least 80% identical, more preferably at least 90% or 95% identical, still more preferably at least 96%, 97%, 98% or 99% identical to the polypeptide encoded by the deposited cDNA, to the polypeptide of Figures 1 A and B (SEQ ID NO:2), and also include portions of such polypeptides with at least 30 amino acids and more preferably at least 50 amino acids. The present invention also includes deletion of the transmembrane region only and retention of at least part of the intracellular domain itself or fusion with at least part of an alternate intracellular domain as described in WO96/04382.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a reference amino acid sequence of a calcitonin receptor polypeptide is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of the calcitonin receptor polypeptide. In other words, to

obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequence shown in Figures 1 A and B (SEQ ID NO:2) or to the amino acid sequence encoded by deposited cDNA clone can be determined conventionally using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711. When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are allowed.

The polypeptide of the present invention could be used as a molecular weight marker on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art.

In another aspect, the invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide described herein. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the

immunogen. On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes. See, for instance, Geysen *et al., Proc. Natl. Acad. Sci. USA 81*:3998-4002 (1983).

As to the selection of peptides or polypeptides bearing an antigenic epitope (i.e., that contain a region of a protein molecule to which an antibody can bind), it is well known in that art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein. See, for instance, Sutcliffe, J. G., Shinnick, T. M., Green, N. and Learner, R.A. (1983). Antibodies that react with predetermined sites on proteins. *Science 219:*660-666. Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are confined neither to immunodominant regions of intact proteins (i.e., immunogenic epitopes) nor to the amino or carboxyl terminals.

Antigenic epitope-bearing peptides and polypeptides of the invention are therefore useful to raise antibodies, including monoclonal antibodies, that bind specifically to a polypeptide of the invention.

Antigenic epitope-bearing peptides and polypeptides of the invention preferably contain a sequence of at least seven, more preferably at least nine and most preferably at least between about 15 to about 30 amino acids contained within the amino acid sequence of a polypeptide of the invention.

Non-limiting examples of antigenic polypeptides or peptides that can be used to generate calcitonin receptor-specific antibodies include:a polypeptide comprising amino acid residues from about 10 to about 20 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 49 to about 60 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 113 to about 123 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 145 to about 154 in SEQ ID NO:2; a polypeptide comprising amino acid residues

from about 189 to about 209 in SEQ ID NO:2; and a polypeptide comprising amino acid residues from about 259 to about 560 in SEQ ID NO:2.

The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means. Houghten, R. A. (1985) General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids. *Proc. Natl. Acad. Sci. USA 82*:5131-5135. This "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Patent No. 4,631,211 to Houghten *et al.* (1986).

As one of skill in the art will appreciate, calcitonin receptor polypeptides of the present invention and the epitope-bearing fragments thereof described above can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life *in vivo*. This has been shown, e.g., for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins (EPA 394,827; Traunecker *et al., Nature 331*:84- 86 (1988)). Fusion proteins that have a disulfide-linked dimeric structure due to the IgG part can also be more efficient in binding and neutralizing other molecules than the monomeric calcitonin receptor protein or protein fragment alone (Fountoulakis *et al., J Biochem 270*:3958-3964 (1995)).

Polynucleotide assays

This invention is also related to the use of the calcitonin receptor polynucleotides to detect complementary polynucleotides such as, for example, as a diagnostic reagent. Detection of a mutated form of calcitonin receptor associated with a dysfunction will provide a diagnostic tool that can add or define a diagnosis of a disease or susceptibility to a disease which results from under-expression or over-expression or altered expression of calcitonin receptor

by assaying the expression level of the gene encoding the calcitonin receptor protein. In addition, individuals carrying mutations in the human calcitonin receptor gene may be detected at the DNA level by a variety of techniques. Nucleic acids for diagnosis may be obtained from a biological sample. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR prior to analysis. PCR (Saiki *et al., Nature, 324*: 163-166 (1986)). RNA or cDNA may also be used in the same ways. As an example, PCR primers complementary to the nucleic acid encoding calcitonin receptor can be used to identify and analyze calcitonin receptor expression and mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to radiolabeled calcitonin receptor RNA or alternatively, radiolabeled calcitonin receptor antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures.

By "assaying the expression level of the gene encoding the calcitonin receptor protein" is intended qualitatively or quantitatively measuring or estimating the level of the calcitonin receptor protein or the level of the mRNA encoding the calcitonin receptor protein in a first biological sample either directly (e.g., by determining or estimating absolute protein level or mRNA level) or relatively (e.g., by comparing to the calcitonin receptor protein level or mRNA level in a second biological sample).

Thus, the invention provides a diagnostic method useful during diagnosis of such disorders, which involves: (a) assaying calcitonin receptor gene expression level in cells or body fluid of an individual; (b) comparing the calcitonin receptor gene expression level with a standard calcitonin receptor gene expression level, whereby an increase or decrease in the assayed calcitonin receptor gene expression level compared to the standard expression level is indicative of such disorders. As will be appreciated in the art, once a standard

calcitonin receptor protein level or mRNA level is known, it can be used repeatedly as a standard for comparison.

By "biological sample" is intended any biological sample obtained from an individual, cell line, tissue culture, or other source which contains calcitonin receptor protein or mRNA. Biological samples include mammalian body fluids (such as sera, plasma, urine, synovial fluid and spinal fluid) which contain calcitonin receptor protein, and ovarian, prostate, heart, placenta, pancreas liver, spleen, lung, breast and umbilical tissue.

Sequence differences between a reference gene and genes having mutations also may be revealed by direct DNA sequencing. In addition, cloned DNA segments may be employed as probes to detect specific DNA segments. The sensitivity of such methods can be greatly enhanced by appropriate use of PCR or another amplification method. For example, a sequencing primer is used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotide or by automatic sequencing procedures with fluorescent-tags.

Genetic testing based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (see, e.g., Myers *et al., Science, 230*: 1242 (1985)).

Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (e.g., Cotton *et al., Proc. Natl. Acad. Sci., USA, 85*: 4397-4401 (1985)).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA

sequencing or the use of restriction enzymes, (e.g., restriction fragment length polymorphisms ("RFLP") and Southern blotting of genomic DNA. Total cellular RNA can be isolated from a biological sample using any suitable technique such as the single-step guanidinium-thiocyanate-phenol-chloroform method described in Chomczynski and Sacchi, *Anal. Biochem. 162:*156-159 (1987). Levels of mRNA encoding the calcitonin receptor protein are then assayed using any appropriate method. These include Northern blot analysis, S1 nuclease mapping, the polymerase chain reaction (PCR), reverse transcription in combination with the polymerase chain reaction (RT-PCR), and reverse transcription in combination with the ligase chain reaction (RT-LCR).

In accordance with a further aspect of the invention, there is provided a process for diagnosing or determining a susceptibility to migrane; infections, such as bacterial, fungal, protozoan and viral infections, particularly infection caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation; dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others. Thus, a mutation in calcitonin receptor indicates a susceptibility to the aforementioned diseases, disorders and conditions and the nucleic acid sequences described above may be employed in an assay for ascertaining such susceptibility. Thus, for example, the assay may be employed to determine a mutation in a human calcitonin receptor protein as herein described, such as a deletion, truncation, insertion, frame shift, etc., with such mutation being indicative of a susceptibility to the aforementioned diseases, disorders and conditions

The invention provides a process for diagnosing to the aforementioned diseases, disorders and conditions comprising determining from a sample derived from a patient an abnormally decreased or increased level of expression

of polynucleotide having the sequence of Figures 1 A and B (SEQ ID NO: 1). Decreased or increased expression of polynucleotide can be measured using any one of the methods well known in the art for the quantation of polynucleotides, such as, for example, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods.

In addition to more conventional gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

Polypeptide assays

The present invention also relates to a diagnostic assays such as quantitative and diagnostic assays for detecting levels of calcitonin receptor protein in cells and tissues, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for detecting over-expression of calcitonin receptor protein compared to normal control tissue samples may be used to detect the presence of a disease/disorder such as infections, including bacterial, fungal, protozoan and viral infections, particularly infection caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation; dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others.

Assaying calcitonin receptor protein levels in a biological sample can occur using any art-known method. For example, calcitonin receptor protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., *et al., J. Cell. Biol. 101:*976-985 (1985); Jalkanen, M., *et al., J. Cell . Biol. 105:*3087-3096 (1987)).

Other antibody-based methods useful for detecting calcitonin receptor protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA).

Suitable labels are known in the art and include enzyme labels, such as glucose oxidase, and radioisotopes, such as iodine (^{125}I, ^{121}I), carbon (^{14}C), salphee (^{35}S), tritium (^{3}H), indium (^{112}In), and technetium (^{99m}Tc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

Antibodies capable of binding to the calcitonin receptor protein antigen may be produced by any art-known method. It will be appreciated that Fab and $F(ab')_2$ and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Alternatively, calcitonin receptor-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or purify the polypeptide of the present invention by attachment of the antibody to a solid support for isolation and/or purification by affinity chromatography.

Thus, among others, antibodies against calcitonin receptor may be employed to inhibit to the aforementioned diseases, disorders and conditions.

Calcitonin Receptor Binding Molecules and Assays

Calcitonin receptor could be used to isolate proteins which interact with it and this interaction could be a target for interference. Inhibitors of protein-protein interactions between calcitonin receptor and other factors could lead to the development of pharmaceutical agents for the modulation of calcitonin receptor activity.

Thus, this invention also provides a method for identification of binding molecules to calcitonin receptor. Genes encoding proteins for binding molecules to calcitonin receptor can be identified by numerous methods known to those of

skill in the art, for example, ligand panning and FACS sorting. Such methods are described in many laboratory manuals such as, for instance, Coligan et al., Current Protocols in Immunology 1 (Rivett, A.J. Biochem. J. 291, 1-10 (1993)): Chapter 5 (1991).

For example, the yeast two-hybrid system provides methods for detecting the interaction between a first test protein and a second test protein, *in vivo*, using reconstitution of the activity of a transcriptional activator. The method is disclosed in U.S. Patent No. 5,283,173; reagents are available from Clontech and Stratagene.

An alternative method is screening of gt11, ZAP (Stratagene) or equivalent cDNA expression libraries with recombinant calcitonin receptor. Recombinant calcitonin receptor protein or fragments thereof are fused to small peptide tags such as FLAG, HSV or GST. The peptide tags can possess convenient phosphorylation sites for a kinase such as heart muscle creatine kinase or they can be biotinylated. Recombinant calcitonin receptor can be phosphorylated with ^{32}[P] or used unlabeled and detected with streptavidin or antibodies against the tags. gt11cDNA expression libraries are made from cells of interest and are incubated with the recombinant calcitonin receptor, washed and cDNA clones isolated which interact with calcitonin receptor. See, e.g., T. Maniatis *et al, infra*..

Another method is the screening of a mammalian expression library in which the cDNAs are cloned into a vector between a mammalian promoter and polyadenylation site and transiently transfected in COS or 293 cells followed by detection of the binding protein. See Sims *et al., Science 241*, 585-589 (1988) and McMahan *et al., EMBO J. 10*, 2821-2832 (1991). In this manner, pools of cDNAs containing the cDNA encoding the binding protein of interest can be selected and the cDNA of interest can be isolated by further subdivision of each pool followed by cycles of transient transfection, binding and autoradiography. Alternatively, the cDNA of interest can be isolated by transfecting the entire cDNA library into mammalian cells and panning the cells on a dish containing

calcitonin receptor bound to the plate. See Seed *et al, Proc. Natl. Acad. Sci. USA 84,* 3365 (1987) and Aruffo *et al., EMBO J. 6,* 3313 (1987). If the binding protein is secreted, its cDNA can be obtained by a similar pooling strategy once a binding or neutralizing assay has been established for assaying supernatants from transiently transfected cells. General methods for screening supernatants are disclosed in Wong *et al., Science 228,* 810-815 (1985).

Another alternative method is immunoaffinity purification well known to those skilled in the art. Yet another alternative method also known to those skilled in the art is screening of peptide libraries for binding partners.

Calcitonin receptor binding partners identified by any of these methods or other methods which would be known to those of ordinary skill in the art as well as those putative binding partners discussed above can be used in the assay method of the invention.

Assays for free calcitonin receptor or binding partner are accomplished by, for example, ELISA or immunoassay using specific antibodies or by incubation of radiolabeled calcitonin receptor with cells or cell membranes followed by centrifugation or filter separation steps. In the presence of test substances which interrupt or inhibit formation of calcitonin receptor/binding partner interaction, an increased amount of free calcitonin receptor or free binding partner will be determined relative to a control lacking the test substance.

Polypeptides of the invention also can be used to assess calcitonin receptor binding capacity of calcitonin receptor binding molecules in cells or in cell-free preparations.

Calcitonin Receptor Agonists and Antagonists

The G-protein coupled receptor family includes dopamine receptors which bind neuroleptic drugs used for treating psychotic and neurological disorders. Other examples of members of this family include adrenergic, endothelin, cAMP, adenosine, muscarinic, acetylcholine, serotonin, histamine, thrombin, kinin, follicle

stimulating hormone, opsins, endothelial differentiation gene-1 receptor, rhodopsins, odorant, cytomegalovirus receptors, etc. Numerous therapeutic agents targeting 7 transmembrane (7 TM) receptors have been successfully introduced onto the market. This indicates that these receptors have an established, proven history as therapeutic targets. G-protein coupled receptors, such as the calcitonin receptor of the present invention, play a role a number of dysfunctions or diseases. Thus, agonists or antagonists of the calcitonin receptor of the present invention could be used therapeutically to prevent, ameliorate or correct dysfunctions or diseases associated with the calcitonin receptor of the present invention.

The calcitonin receptor of the present invention may be employed in a process for screening for compounds which activate (agonists) or inhibit activation (antagonists) of the receptor polypeptide of the present invention .

In general, such screening procedures involve providing appropriate cells which express the receptor polypeptide of the present invention on the surface thereof. Such cells include cells from mammals, yeast, drosophila or *E. coli*. In particular, a polynucleotide encoding the receptor of the present invention is employed to transfect cells to thereby express the calcitonin receptor. The expressed receptor is then contacted with a test compound to observe binding, stimulation or inhibition of a functional response.

One such screening procedure involves the use of melanophores which are transfected to express the calcitonin receptor of the present invention. Such a screening technique is described in PCT WO 92/01810 published February 6, 1992.

Thus, for example, such assay may be employed for screening for a compound which inhibits activation of the receptor polypeptide of the present invention by contacting the melanophore cells which encode the receptor with both the receptor ligand and a compound to be screened. Inhibition of the signal generated by the ligand indicates that a compound is a potential antagonist for the receptor, i.e., inhibits activation of the receptor.

The screen may be employed for determining a compound which activates the receptor by contacting such cells with compounds to be screened and determining whether such compound generates a signal, i.e., activates the receptor.

Other screening techniques include the use of cells which express the calcitonin receptor (for example, transfected CHO cells) in a system which measures extracellular pH changes caused by receptor activation, for example, as described in *Science, 246*: 181-296 (October 1989). For example, compounds may be contacted with a cell which expresses the receptor polypeptide of the present invention and a second messenger response, e.g. signal transduction or pH changes, may be measured to determine whether the potential compound activates or inhibits the receptor.

Another method involves screening for compounds which inhibit activation of the receptor polypeptide of the present invention antagonists by determining inhibition of binding of labeled ligand to cells which have the receptor on the surface thereof. Such a method involves transfecting a eukaryotic cell with DNA encoding the calcitonin receptor such that the cell expresses the receptor on its surface and contacting the cell with a compound in the presence of a labeled form of a known ligand. The ligand can be labeled, e.g., by radioactivity. The amount of labeled ligand bound to the receptors is measured, e.g., by measuring radioactivity associated with transfected cells or membrane from these cells. If the compound binds to the receptor as determined by a reduction of labeled ligand which binds to the receptors, the binding of labeled ligand to the receptor is inhibited.

Another method involves screening for calcitonin receptor inhibitors by determining inhibition or stimulation of calcitonin receptor-mediated cAMP and/or adenylate cyclase accumulation. Such a method involves transfecting a eukaryotic cell with calcitonin receptor to express the receptor on the cell surface. The cell is then exposed to potential antagonists in the presence of calcitonin receptor. The amount of cAMP accumulation is then measured. If the potential antagonist binds the receptor, and thus inhibits calcitonin receptor binding, the levels of calcitonin

receptor-mediated cAMP, or adenylate cyclase, activity will be reduced or increased.

The present invention also provides a method for determining whether a ligand not known to be capable of binding to a calcitonin receptor can bind to such receptor which comprises contacting a mammalian cell which expresses a calcitonin receptor with the ligand under conditions permitting binding of ligands to the calcitonin receptor, detecting the presence of a ligand which binds to the receptor and thereby determining whether the ligand binds to the calcitonin receptor. The systems hereinabove described for determining agonists and/or antagonists may also be employed for determining ligands which bind to the receptor.

Examples of potential calcitonin receptor receptor antagonists are an antibody, or in some cases an oligonucleotide, which binds to the receptor but does not elicit a second messenger response such that the activity of the receptor is prevented.

Potential antagonists also include proteins which are closely related to the ligand of the calcitonin receptor, i.e. a fragment of the ligand, which have lost biological function and when binding to the calcitonin receptor, elicit no response.

A potential antagonist also includes an antisense construct prepared through the use of antisense technology. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes for the polypeptides of the present invention, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix -see Lee *et al., Nucl. Acids Res., 6*:3073 (1979); Cooney *et al, Science, 241*:456 (1988); and Dervan *et al., Science, 251*: 1360 (1991)), thereby preventing transcription and the production of calcitonin receptor. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the calcitonin receptor (antisense - Okano,

J. Neurochem., *56*:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the calcitonin receptor.

Another potential antagonist is a small molecule which binds to the calcitonin receptor, making it inaccessible to ligands such that normal biological activity is prevented. Examples of small molecules include but are not limited to small peptides or peptide-like molecules.

Potential antagonists also include a soluble form of a calcitonin receptor, e.g. a fragment of the receptor, which binds to the ligand and prevents the ligand from interacting with membrane bound calcitonin receptors.

Calcitonin receptor are ubiquitous in the mammalian host and are responsible for many biological functions, including many pathologies. Accordingly, it is desirous to find compounds and drugs which stimulate the calcitonin receptor on the one hand and which can inhibit the function of a calcitonin receptor on the other hand.

In general, agonists or antagonists for calcitonin receptor are employed for therapeutic and prophylactic purposes for such diseases or disorders as infections, such as bacterial, fungal, protozoan and viral infections, particularly infection caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation; dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others.

Modes of administration

It will be appreciated that conditions caused by a decrease in the standard or normal level of calcitonin receptor activity in an individual, can be treated by administration of agonists of the calcitonin receptor protein. The invention further provides a method of treating an individual in need of an increased level of calcitonin receptor activity comprising administering to such an individual a pharmaceutical composition comprising an effective amount of an agonist of the calcitonin receptor of the invention, effective to increase the calcitonin receptor activity level in such an individual.

As a general proposition, the total pharmaceutically effective amount of calcitonin receptor agonist or antagonist administered parenterally per dose will be in the range of about 1 μg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the calcitonin receptor polypeptide is typically administered at a dose rate of about 1 μg/kg/hour to about 50 μg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the calcitonin receptor agonist or antagonist may be administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by powders, ointments, drops or transdermal patch), bucally, or as an oral or nasal spray. By "pharmaceutically acceptable carrier" is meant a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

Chromosome Assays

The nucleic acid molecules of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

In certain preferred embodiments in this regard, the cDNA herein disclosed is used to clone genomic DNA of a calcitonin receptor protein gene. This can be accomplished using a variety of well known techniques and libraries, which generally are available commercially. The genomic DNA then is used for *in situ* chromosome mapping using well known techniques for this purpose. Typically, in accordance with routine procedures for chromosome mapping, some trial and error may be necessary to identify a genomic probe that gives a good *in situ* hybridization signal.

In addition, in some cases, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the 3′ untranslated region of the gene is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes.

Fluorescence *in situ* hybridization ("FISH") of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with probes from the cDNA as short as 50 or 60 bp. For a review of this technique, see Verma *et al.*, *Human Chromosomes: A Manual Of Basic Techniques*, Pergamon Press, New York (1988).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, *Mendelian Inheritance In Man*, available on-line through Johns Hopkins University, Welch Medical Library. The relationship between genes and diseases that have been mapped to the same

chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1: Expression and Purification of Calcitonin Receptor in E. coli

The DNA sequence encoding the calcitonin receptor protein in the deposited cDNA clone is amplified using PCR oligonucleotide primers specific to the amino terminal sequences of the calcitonin receptor protein and to vector sequences 3′ to the gene. Additional nucleotides containing restriction sites to facilitate cloning are added to the 5′ and 3′ sequences respectively.

The 5′ oligonucleotide primer has the sequence 5′ GCCAGGTCGACCACATGACGAGCTTCGCT 3′ (SEQ ID NO:3) containing the underlined Sal I restriction site, which encodes 15 nucleotides of the calcitonin receptor protein coding sequence in SEQ ID NO:1.

The 3′ primer has the sequence 5′ CGGCTCTAGAGTTAATGCACGAAGTTAAAGAAG 3′ (SEQ ID NO:4) containing the underlined Xba I restriction site followed by nucleotides complementary to nucleotides immediately after the calcitonin receptor protein coding sequence in SEQ ID NO:1.

The restriction sites are convenient to restriction enzyme sites in the bacterial expression vector pD10 (pQE9), which are used for bacterial expression in these examples. (Qiagen, Inc. 9259 Eton Avenue, Chatsworth, CA, 91311). [pD10]pQE9

encodes ampicillin antibiotic resistance ("Amp") and contains a bacterial origin of replication ("ori"), an IPTG inducible promoter, a ribosome binding site ("RBS"), a 6-His tag and restriction enzyme sites.

The amplified calcitonin receptor DNA and the vector pQE9 both are digested with SalI and XbaI and the digested DNAs are then ligated together. Insertion of the calcitonin receptor protein DNA into the restricted pQE9 vector places the calcitonin receptor protein coding region downstream of and operably linked to the vector's IPTG-inducible promoter and in-frame with an initiating AUG appropriately positioned for translation of calcitonin receptor protein.

The ligation mixture is transformed into competent *E. coli* cells using standard procedures. Such procedures are described in Sambrook *et al.*, Molecular Cloning: a Laboratory Manual, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). *E. coli* strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses lac repressor and confers kanamycin resistance ("Kan"), is used in carrying out the illustrative example described herein. This strain, which is only one of many that are suitable for expressing calcitonin receptor protein, is available commercially from Qiagen.

Transformants are identified by their ability to grow on LB plates in the presence of ampicillin and kanamycin. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 µg/ml) and kanamycin (25 µg/ml).

The O/N culture is used to inoculate a large culture, at a dilution of approximately 1:100 to 1:250. The cells are grown to an optical density at 600 nm ("OD600") of between 0.4 and 0.6. Isopropyl-B-D-thiogalactopyranoside ("IPTG") is then added to a final concentration of 1 mM to induce transcription from *lac* repressor sensitive promoters, by inactivating the *lacI* repressor. Cells subsequently are incubated further for 3 to 4 hours. Cells then are harvested by centrifugation and disrupted, by standard methods. Inclusion bodies are purified from the disrupted cells

using routine collection techniques, and protein is solubilized from the inclusion bodies into 8M urea. The 8M urea solution containing the solubilized protein is passed over a PD-10 column in 2X phosphate-buffered saline ("PBS"), thereby removing the urea, exchanging the buffer and refolding the protein. The protein is purified by a further step of chromatography to remove endotoxin. Then, it is sterile filtered. The sterile filtered protein preparation is stored in 2X PBS at a concentration of 95 μ/ml.

Example 2: Cloning and Expression of Calcitonin Receptor protein in a Baculovirus Expression System

The cDNA sequence encoding the full length calcitonin receptor protein in the deposited clone is amplified using PCR oligonucleotide primers corresponding to the 5′ and 3′ sequences of the gene:

The 5′ primer has the sequence 5′GTCATCTAGAAGTCCGCCATCATGACGAGCAGCTTCGCTGTGCTC 3′ (SEQ ID NO:5) containing the underlined Xba I restriction enzyme site followed by 21 bases of the sequence of calcitonin receptor protein in SEQ ID NO:1. Inserted into an expression vector, as described below, the 5′ end of the amplified fragment encoding calcitonin receptor provides an efficient signal peptide. An efficient signal for initiation of translation in eukaryotic cells, as described by Kozak, M., *J. Mol. Biol. 196*:947-950 (1987) is appropriately located in the vector portion of the construct.

The 3′ primer has the sequence 5′ CGGCTCTAGAGTTAATGCACGAAGTTAAAGAAG 3′ (SEQ ID NO:6) containing the underlined Xba I restriction site followed by nucleotides complementary to the last 19 nucleotides of the calcitonin receptor coding sequence set out in SEQ ID NO:1.

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment

then is digested with XbaI and again is purified on a 1% agarose gel. This fragment is designated herein F2.

The vector pA2-GP is used to express the calcitonin receptor protein in the baculovirus expression system, using standard methods, as described in Summers *et al.*, A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures, Texas Agricultural Experimental Station Bulletin No. 1555 (1987). This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites. The signal peptide of AcMNPV gp67, including the N-terminal methionine, is located just upstream of a BamHI site. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For an easy selection of recombinant virus the beta-galactosidase gene from *E. coli* is inserted in the same orientation as the polyhedrin promoter and is followed by the polyadenylation signal of the polyhedrin gene. The polyhedrin sequences are flanked at both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate viable virus that express the cloned polynucleotide.

Many other baculovirus vectors could be used in place of pA2-GP, such as pAc373, pVL941 and pAcIM1 provided, as those of skill readily will appreciate, that construction provides appropriately located signals for transcription, translation, trafficking and the like, such as an in-frame AUG and a signal peptide, as required. Such vectors are described in Luckow *et al.*, *Virology 170*: 31-39, among others.

The plasmid is digested with the restriction enzyme XbaI and then is dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.). This vector DNA is designated herein "V".

Fragment F2 and the dephosphorylated plasmid V2 are ligated together with T4 DNA ligase. *E. coli* HB101 cells are transformed with ligation mix and spread on culture plates. Bacteria are identified that contain the plasmid with the human calcitonin receptor gene by digesting DNA from individual colonies using XbaI and

then analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing. This plasmid is designated herein pBaccalcitonin receptor.

5 μg of the plasmid pBaccalcitonin receptor is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA.), using the lipofection method described by Felgner *et al.*, Proc. Natl. Acad. Sci. USA 84: 7413-7417 (1987). 1μg of BaculoGold™ virus DNA and 5 μg of the plasmid pBaccalcitonin receptor are mixed in a sterile well of a microtiter plate containing 50 μl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is rocked back and forth to mix the newly added solution. The plate is then incubated for 5 hours at 27°C. After 5 hours the transfection solution is removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. The plate is put back into an incubator and cultivation is continued at 27°C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, cited above. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10).

Four days after serial dilution, the virus is added to the cells. After appropriate incubation, blue stained plaques are picked with the tip of an Eppendorf pipette. The agar containing the recombinant viruses is then resuspended in an Eppendorf tube containing 200 μl of Grace's medium. The agar is removed by a brief centrifugation and the supernatant containing the recombinant baculovirus is used to

infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4°C. A clone containing properly inserted hESSB I, II and III is identified by DNA analysis including restriction mapping and sequencing. This is designated herein as V-calictonin receptor.

Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus V-calcitonin receptor at a multiplicity of infection ("MOI") of about 2 (about 1 to about 3). Six hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Gaithersburg). 42 hours later, 5 μCi of ^{35}S-methionine and 5 μCi ^{35}S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then they are harvested by centrifugation, lysed and the labeled proteins are visualized by SDS-PAGE and autoradiography.

Example 3: Cloning and Expression in Mammalian Cells

Most of the vectors used for the transient expression of the calcitonin receptor protein gene sequence in mammalian cells should carry the SV40 origin of replication. This allows the replication of the vector to high copy numbers in cells (e.g., COS cells) which express the T antigen required for the initiation of viral DNA synthesis. Any other mammalian cell line can also be utilized for this purpose.

A typical mammalian expression vector contains the promoter element, which mediates the initiation of transcription of mRNA, the protein coding sequence, and signals required for the termination of trancription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription can be achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular signals can also be used

(e.g., human actin promoter). Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) and pBC12MI (ATCC 67109). Mammalian host cells that could be used include, human Hela, 283, H9 and Jurkart cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, African green monkey cells, quail QC1-3 cells, mouse L cells and Chinese hamster ovary cells.

Alternatively, the gene can be expressed in stable cell lines that contain the gene integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) is a useful marker to develop cell lines that carry several hundred or even several thousand copies of the gene of interest. Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy *et al., Biochem J. 227*:277-279 (1991); Bebbington *et al., Bio/Technology 10*:169-175 (1992)). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) cells are often used for the production of proteins.

The expression vectors pC1 and pC4 contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen *et al., Molecular and Cellular Biology*, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart *et al., Cell 41*:521-530 (1985)). Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors contain in addition the 3′ intron, the polyadenylation and termination signal of the rat preproinsulin gene.

Example 3(a): Cloning and Expression in COS Cells

The expression plasmid, pCalcitonin ReceptorHA, is made by cloning a cDNA encoding calcitonin receptor into the expression vector pcDNAI/Amp (which can be obtained from Invitrogen, Inc.).

The expression vector pcDNAI/amp contains: (1) an *E.coli* origin of replication effective for propagation in *E. coli* and other prokaryotic cells; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40 origin of replication for propagation in eukaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron, and a polyadenylation signal arranged so that a cDNA conveniently can be placed under expression control of the CMV promoter and operably linked to the SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker.

A DNA fragment encoding the calcitonin receptor protein and an HA tag fused in frame to its 3′ end is cloned into the polylinker region of the vector so that recombinant protein expression is directed by the CMV promoter. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein described by Wilson *et al., Cell 37:* 767 (1984). The fusion of the HA tag to the target protein allows easy detection of the recombinant protein with an antibody that recognizes the HA epitope.

The plasmid construction strategy is as follows. The calcitonin receptor cDNA of the deposited clone is amplified using primers that contain convenient restriction sites, much as described above regarding the construction of expression vectors for expression of calcitonin receptor in *E. coli*. To facilitate detection, purification and characterization of the expressed calcitonin receptor, one of the primers contains a hemagglutinin tag ("HA tag") as described above.

Suitable primers include the following, which are used in this example. The 5′ primer, containing the underlined Xba I site, an AUG start codon and 21 bases of the 5′ coding region has the following sequence:

5'GTCATCTAGAAGTCCGCCATCATGACGAGCTTCGCTG TGCTC 3' (SEQ ID NO:7).

The 3' primer, containing the underlined Xba I site, a stop codon, 9 codons thereafter forming the hemagglutinin HA tag, and 19 bases of 3' coding sequence (at the 3' end) has the following sequence:

5'CGGCTCTAGAGTTAAGCGTAGTCTGGGACGTCGTATG GGTAATGCACGAAGGTTAAAGAAG 3' (SEQ ID NO:8).

The following 3' primer, containing the underlined Xba I site, a stop codon, and 19 bases of 3' coding sequence (at the 3' end) can be used to express a protein lacking the HA tag. This primer has the following sequence:

5' CGGCTCTAGAGTTAATGCACGAAGTTAAAGAAG 3' (SEQ ID NO:9).

The PCR amplified DNA fragment and the vector, pcDNAI/Amp, are digested with HindIII and XhoI and then ligated. The ligation mixture is transformed into *E. coli* strain SURE (available from Stratagene Cloning Systems, 11099 North Torrey Pines Road, La Jolla, CA 92037), and the transformed culture is plated on ampicillin media plates which then are incubated to allow growth of ampicillin resistant colonies. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis and gel sizing for the presence of the calcitonin receptor-encoding fragment.

For expression of recombinant calcitonin receptor, COS cells are transfected with an expression vector, as described above, using DEAE-DEXTRAN, as described, for instance, in Sambrook *et al.*, Molecular Cloning: a Laboratory Manual, Cold Spring Laboratory Press, Cold Spring Harbor, New York (1989). Cells are incubated under conditions for expression of calcitonin receptor by the vector.

Expression of the calcitonin receptor HA fusion protein is detected by radiolabelling and immunoprecipitation, using methods described in, for example Harlow *et al.*, Antibodies: A Laboratory Manual, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988). To this end, two days after transfection, the cells are labeled by incubation in media containing ^{35}S-cysteine for

8 hours. The cells and the media are collected, and the cells are washed and the lysed with detergent-containing RIPA buffer: 150 mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50 mM TRIS, pH 7.5, as described by Wilson *et al.* cited above. Proteins are precipitated from the cell lysate and from the culture media using an HA-specific monoclonal antibody. The precipitated proteins then are analyzed by SDS-PAGE gels and autoradiography. An expression product of the expected size is seen in the cell lysate, which is not seen in negative controls.

Example 3(b): Cloning and Expression in CHO Cells

The vector pC4 is used for the expression of calcitonin receptor protein. Plasmid pC4 is a derivative of the plasmid pSV2-dhfr [ATCC Accession No. 37146]. Both plasmids contain the mouse DHFR gene under control of the SV40 early promoter. Chinese hamster ovary- or other cells lacking dihydrofolate activity that are transfected with these plasmids can be selected by growing the cells in a selective medium (alpha minus MEM, Life Technologies) supplemented with the chemotherapeutic agent methotrexate. The amplification of the DHFR genes in cells resistant to methotrexate (MTX) has been well documented (see, e.g., Alt, F.W., Kellems, R.M., Bertino, J.R., and Schimke, R.T., 1978, J. Biol. Chem. 253:1357-1370, Hamlin, J.L. and Ma, C. 1990, Biochem. et Biophys. Acta, 1097:107-143, Page, M.J. and Sydenham, M.A. 1991, Biotechnology Vol. 9:64-68). Cells grown in increasing concentrations of MTX develop resistance to the drug by overproducing the target enzyme, DHFR, as a result of amplification of the DHFR gene. If a second gene is linked to the DHFR gene it is usually co-amplified and over-expressed. It is state of the art to develop cell lines carrying more than 1,000 copies of the genes. Subsequently, when the methotrexate is withdrawn, cell lines contain the amplified gene integrated into the chromosome(s).

Plasmid pC4 contains for expressing the gene of interest the strong promoter of the long terminal repeat (LTR) of the Rous Sarcoma Virus (Cullen, *et al., Molecular and Cellular Biology,* March 1985:438-447) plus a fragment isolated from

the enhancer of the immediate early gene of human cytomegalovirus (CMV) (Boshart *et al., Cell 41*:521-530 (1985)). Downstream of the promoter are Bam HI, XbaI, and Asp718 restriction enzyme cleavage sites that allow integration of the genes. Behind these cloning sites the plasmid contains the 3' intron and polyadenylation site of the rat preproinsulin gene. Other high efficiency promoters can also be used for the expression, e.g., the human β-actin promoter, the SV40 early or late promoters or the long terminal repeats from other retroviruses, e.g., HIV and HTLVI. Clontech's Tet-Off and Tet-On gene expression systems and similar systems can be used to express the Calcitonin Receptor in a regulated way in mammalian cells (Gossen, M., & Bujard, H. 1992, *Proc. Natl. Acad. Sci. USA 89*: 5547-5551). For the polyadenylation of the mRNA other signals, e.g., from the human growth hormone or globin genes can be used as well. Stable cell lines carrying a gene of interest integrated into the chromosomes can also be selected upon co-transfection with a selectable marker such as gpt, G418 or hygromycin. It is advantageous to use more than one selectable marker in the beginning, e.g., G418 plus methotrexate.

Stable cell lines carrying a gene of interest integrated into the chromosomes can also be selected upon co-transfection with a selectable marker such as gpt, G418 or hygromycin. It is advantageous to use more than one selectable marker in the beginning, e.g., G418 plus methotrexate.

The plasmid pC4 is digested with the restriction enzyme Xba I and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

The DNA sequence encoding calcitonin receptor, ATCC No. 97730, is amplified using PCR oligonucleotide primers corresponding to the 5′ and 3′ sequences of the gene:

The 5′ primer has the sequence 5′ GTCATCTAGAAGTCCGCCATCATGACGAGCGGCGCTGTGCTC 3′ (SEQ ID NO:10) containing the underlined Xba I restriction enzyme site followed by 21 bases of the sequence of calcitonin receptor of SEQ ID NO:1. Inserted into an expression vector, as described below, the 5′ end of the amplified fragment encoding

human calcitonin receptor provides an efficient signal peptide. An efficient signal for initiation of translation in eukaryotic cells, as described by Kozak, M., J. Mol. Biol. 196:947-950 (1987) is appropriately located in the vector portion of the construct.

The 3′ primer, containing the underlined Xba I site, a stop codon, 9 codons thereafter forming the hemagglutinin HA tag, and 19 bases complementary to the 3′ coding sequence (at the 3′end) has the following sequence:

5′CGGCTCTAGAGTTAAGCGTAGTCTGGGACGTCGTATGGGTAATGCA CGAAGGTTAAAGAAG 3′ (SEQ ID NO:11).

The following 3′ primer, containing the underlined Xba I site, a stop codon, and 19 bases of 3′ coding sequence (at the 3′ end) can be used to express a protein lacking the HA tag. This primer has the following sequence:

5′ CGGCTCTAGAGTTAATGCACGAAGTTAAAGAAG 3′ (SEQ ID NO:12).

The amplified fragment is digested with the endonuclease Xba I and then purified again on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC4 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene are used for transfection. 5 μg of the expression plasmid pC4 is cotransfected with 0.5 μg of the plasmid pSV2-neo using lipofectin (Felgner *et al., supra).* The plasmid pSV2neo contains a dominant selectable marker, the neo gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of metothrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher

concentrations of methotrexate (1 μM, 2 μM, 5 μM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 μM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reverse phase HPLC analysis.

Example 4: Tissue distribution of Calcitonin Receptor protein expression

Northern blot analysis is carried out to examine calcitonin receptor gene expression in human tissues, using methods described by, among others, Sambrook *et al.*, cited above. A cDNA probe containing the entire nucleotide sequence of the calcitonin receptor protein (SEQ ID NO:1) is labeled with ^{32}P using the *redi*prime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labelling, the probe is purified using a CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labelled probe is then used to examine various human tissues for calcitonin receptor mRNA.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM)are obtained from Clontech and are examined with labelled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and films developed according to standard procedures.

Similarly, Northern blots of fetal tissue or cancerous tissues are also examined with labelled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and films developed according to standard procedures.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of all publications (including patents, patent applications, journal articles, laboratory manuals, books, or other documents) cited herein are hereby incorporated by reference.

SEQUENCE LISTING

(1) GENERAL INFORMATION:
(i) APPLICANT: HUMAN GENOME SCIENCES, INC.
9410 KEY WEST AVENUE
ROCKVILLE, MD 20850
UNITED STATES OF AMERICA
INVENTORS: SOPPET, DANIEL R.
LI, YI
RUBEN, STEVEN M.

(ii) TITLE OF INVENTION: CALCITONIN RECEPTOR

(iii) NUMBER OF SEQUENCES: 26

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
(B) STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
(C) CITY: WASHINGTON
(D) STATE: D.C.
(E) COUNTRY: US
(F) ZIP: 20005-3934

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Herewith
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/030,934
(B) FILING DATE: 15NOV1996

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: STEFFE, ERIC K.
(B) REGISTRATION NUMBER: 36,688
(C) REFERENCE/DOCKET NUMBER: 1488.066PC01/EKS/KMT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (202) 371-2600
(B) TELEFAX: (202) 371-2540

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2603 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 652..2355

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CCTTTCCCAC CAGCCGCAGA GAGCCAGGAT GGACGTTCCT CGGACGGACG GTTTTCCTGC      60

TTGGGAATGT TCCTGGGCTG TGAGATCCAC TCTTCTGGGC AGGTGGTTAG CACCTAACGT     120

TTTTCCCTCA CTTCCCCCCA AATTCTTAAG TCCTTTGGTC CATTTCACTG CTCGGACCTT     180

GAGACAACAG TCATTCTGCC TGAGTCTGTC TTCAGAGAGA CGCCCCCCGT GGTCAGGCCC     240

GCAGGCCCCG GAGAGGCCCA GGAGCCAGAG GAGCTGGCAC GGCGACAGCG ACGGCACCCG     300

GAGCTGAGCC AGGGTGAGGC TGTGGCCAGC GTCATCATCT ACCGCACCCT GGCCGGGCTA     360

CTGCCTCATA ACTATGACCC TGACAAGCGC AGCTTGAGAG TCCCCAAACG CCCGATCATC     420

AACACACCCG TGGTGAGCAT CAGCGTCCAT GATGATGAGG AGCTTCTGCC CCGGGCCCTG     480

GACAAACCCG TCACGGTGCA GTTCCGCCTG CTGGAGACAG AGGAGCGGAC CAAGCCCATC     540

TGTGTCTTCT GGAACCATTC AATCCTGGTC AGTGGCACAG GTGGCTGGTC GGCCAGAGGC     600

TGTGAAGTCG TCTTCCGCAA TGAGAGCCAC GTCAGCTGCC AGTTCAACCA C ATG ACG      657
                                                          Met Thr
                                                            1

AGC TTC GCT GTG CTC ATG GAC GTT TCT CGG CGG GAG AAT GGG GAG ATC      705
Ser Phe Ala Val Leu Met Asp Val Ser Arg Arg Glu Asn Gly Glu Ile
          5                  10                  15

CTG CCA CTG AAG ACA CTG ACA TAC GTG GCT CTA GGT GTC GCC TTG GCT      753
Leu Pro Leu Lys Thr Leu Thr Tyr Val Ala Leu Gly Val Ala Leu Ala
     20                  25                  30

GCC CTT CTG CTC ACC TTC TTC TTC CTC ACT CTC TTG CGT ATC CTG CGC      801
Ala Leu Leu Leu Thr Phe Phe Phe Leu Thr Leu Leu Arg Ile Leu Arg
 35                  40                  45                  50

TCC AAC CAA CAC GGC ATC CGA CGT AAC CTG ACA GCT GCC CTG GGC CTG      849
Ser Asn Gln His Gly Ile Arg Arg Asn Leu Thr Ala Ala Leu Gly Leu
                55                  60                  65

GCT CAG CTG GTC TTC CTC CTG GGA ATC AAC CAG GCT GAC CTC CCT TTT      897
Ala Gln Leu Val Phe Leu Leu Gly Ile Asn Gln Ala Asp Leu Pro Phe
            70                  75                  80

GCC TGC ACA GTC ATT GCC ATC CTG CTG CAC TTC CTG TAC CTC TGC ACC      945
Ala Cys Thr Val Ile Ala Ile Leu Leu His Phe Leu Tyr Leu Cys Thr
        85                  90                  95

TTT TCC TGG GCT CTG CTG GAG GCC TTG CAC CTG TAC CGG GCA CTC ACT      993
Phe Ser Trp Ala Leu Leu Glu Ala Leu His Leu Tyr Arg Ala Leu Thr
    100                 105                 110

GAG GTG CGC GAT GTC AAC ACC GGC CCC ATG CGC TTC TAC TAC ATG CTG     1041
Glu Val Arg Asp Val Asn Thr Gly Pro Met Arg Phe Tyr Tyr Met Leu
115                 120                 125                 130

GGC TGG GGC GTG CCT GCC TTC ATC ACA GGG CTA GCC GTG GGC CTG GAC     1089
Gly Trp Gly Val Pro Ala Phe Ile Thr Gly Leu Ala Val Gly Leu Asp
                135                 140                 145
```

```
CCC GAG GGC TAC GGG AAC CCT GAC TTC TGC TGG CTC TCC ATC TAT GAC      1137
Pro Glu Gly Tyr Gly Asn Pro Asp Phe Cys Trp Leu Ser Ile Tyr Asp
            150                 155                 160

ACG CTC ATC TGG AGT TTT GCT GGC CCG GTG GCC TTT GCC GTC TCG ATG      1185
Thr Leu Ile Trp Ser Phe Ala Gly Pro Val Ala Phe Ala Val Ser Met
        165                 170                 175

AGT GTC TTC CTG TAC ATC CTG GCG GCC CGG GCC TCC TGT GCT GCC CAG      1233
Ser Val Phe Leu Tyr Ile Leu Ala Ala Arg Ala Ser Cys Ala Ala Gln
    180                 185                 190

CGG CAG GGC TTT GAG AAG AAA GGT CCT GTC TCG GGC CTG CAG CCC TCC      1281
Arg Gln Gly Phe Glu Lys Lys Gly Pro Val Ser Gly Leu Gln Pro Ser
195                 200                 205                 210

TTC GCC GTC CTC CTG CTG CTG AGC GCC ACG TGG CTG CTG GCA CTG CTC      1329
Phe Ala Val Leu Leu Leu Leu Ser Ala Thr Trp Leu Leu Ala Leu Leu
                215                 220                 225

TCT GTC AAC AGC GAC ACC CTC CTC TTC CAC TAC CTC TTT GCT ACC TGC      1377
Ser Val Asn Ser Asp Thr Leu Leu Phe His Tyr Leu Phe Ala Thr Cys
            230                 235                 240

AAT TGC ATC CAG GGC CCC TTC ATC TTC CTC TCC TAT GTG GTG CTT AGC      1425
Asn Cys Ile Gln Gly Pro Phe Ile Phe Leu Ser Tyr Val Val Leu Ser
        245                 250                 255

AAG GAG GTC CGG AAA GCA CTC AAG CTT GCC TGC AGC CGC AAG CCC AGC      1473
Lys Glu Val Arg Lys Ala Leu Lys Leu Ala Cys Ser Arg Lys Pro Ser
    260                 265                 270

CCT GAC CCT GCT CTG ACC ACC AAG TCC ACC CTG ACC TCG TCC TAC AAC      1521
Pro Asp Pro Ala Leu Thr Thr Lys Ser Thr Leu Thr Ser Ser Tyr Asn
275                 280                 285                 290

TGC CCC AGC CCC TAC GCA GAT GGG CGG CTG TAC CAG CCC TAC GGA GAC      1569
Cys Pro Ser Pro Tyr Ala Asp Gly Arg Leu Tyr Gln Pro Tyr Gly Asp
                295                 300                 305

TCG GCC GGC TCT CTG CAC AGC ACC AGT CGC TCG GGC AAG AGT CAG CCC      1617
Ser Ala Gly Ser Leu His Ser Thr Ser Arg Ser Gly Lys Ser Gln Pro
            310                 315                 320

AGC TAC ATC CCC TTC TTG CTG AGG GAG GAG TCC GCA CTG AAC CCT GGC      1665
Ser Tyr Ile Pro Phe Leu Leu Arg Glu Glu Ser Ala Leu Asn Pro Gly
        325                 330                 335

CAA GGG CCC CCT GGC CTG GGG GAT CCA GGC AGC CTG TTC CTG GAA GGT      1713
Gln Gly Pro Pro Gly Leu Gly Asp Pro Gly Ser Leu Phe Leu Glu Gly
    340                 345                 350

CAA GAC CAG CAG CAT GAT CCT GAC ACG GAC TCC GAC AGT GAC CTG TCC      1761
Gln Asp Gln Gln His Asp Pro Asp Thr Asp Ser Asp Ser Asp Leu Ser
355                 360                 365                 370

TTA GAA GAC GAC CAG AGT GGC TCC TAT GCC TCT ACC CAC TCA TCA GAC      1809
Leu Glu Asp Asp Gln Ser Gly Ser Tyr Ala Ser Thr His Ser Ser Asp
                375                 380                 385

AGT GAG GAG GAA GAA GAG GAG GAG GAA GAG GAG GCC GCC TTC CCT GGA      1857
Ser Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Ala Ala Phe Pro Gly
            390                 395                 400
```

```
GAG CAG GGC TGG GAT AGC CTG CTG GGG CCT GGA GCA GAG AGA CTG CCC      1905
Glu Gln Gly Trp Asp Ser Leu Leu Gly Pro Gly Ala Glu Arg Leu Pro
        405                 410                 415

CTG CAC AGT ACT CCC AAG GAT GGG GGC CCA GGG CCT GGC AAG GCC CCC      1953
Leu His Ser Thr Pro Lys Asp Gly Gly Pro Gly Pro Gly Lys Ala Pro
    420                 425                 430

TGG CCA GGA GAC TTT GGG ACC ACA GCA AAA GAG AGT AGT GGC AAC GGG      2001
Trp Pro Gly Asp Phe Gly Thr Thr Ala Lys Glu Ser Ser Gly Asn Gly
435                 440                 445                 450

GCC CCT GAG GAG CGG CTG CGG GAG AAT GGA GAT GCC CTG TCT CGA GAG      2049
Ala Pro Glu Glu Arg Leu Arg Glu Asn Gly Asp Ala Leu Ser Arg Glu
                455                 460                 465

GGG TCC CTA GGC CCC CTT CCA GGC TCT TCT GCC CAG CCT CAC AAA GGC      2097
Gly Ser Leu Gly Pro Leu Pro Gly Ser Ser Ala Gln Pro His Lys Gly
            470                 475                 480

ATC CTT AAG AAG AAG TGT CTG CCC ACC ATC AGC GAG AAG AGC AGC CTC      2145
Ile Leu Lys Lys Lys Cys Leu Pro Thr Ile Ser Glu Lys Ser Ser Leu
        485                 490                 495

CTG CGG CTC CCC CTG GAG CAA TGC ACA GGG TCT TCC CGG GGC TCC TCC      2193
Leu Arg Leu Pro Leu Glu Gln Cys Thr Gly Ser Ser Arg Gly Ser Ser
    500                 505                 510

GCT AGT GAG GGC AGC CGG GGC GGC CCC CCT CCC CGC CCA CCG CCC CGG      2241
Ala Ser Glu Gly Ser Arg Gly Gly Pro Pro Pro Arg Pro Pro Pro Arg
515                 520                 525                 530

CAG AGC CTC CAG GAG CAG CTG AAC GGG GTC ATG CCC ATC GCC ATG AGC      2289
Gln Ser Leu Gln Glu Gln Leu Asn Gly Val Met Pro Ile Ala Met Ser
                535                 540                 545

ATC AAG GCA GGC ACG GTG GAT GAG GAC TCG TCA GGC TCC GAA TTT CTC      2337
Ile Lys Ala Gly Thr Val Asp Glu Asp Ser Ser Gly Ser Glu Phe Leu
            550                 555                 560

TTC TTT AAC TTC GTG CAT TAACCCTGGG CCGTGGTTCC TACGCCCGAG             2385
Phe Phe Asn Phe Val His
        565

GCTCCCTTCC CTTCCCCAGC CGCACTCATG CCCTGCTCCT GTCTTGTGCT TTATCCTGCC    2445

CCGCTCCCCA TCGCCTGCCC GCAGCAGCGA CGAAACGTCC ATCTGAGGAG CCTGGGCCTT    2505

GCCGGGAGGG GTACTCACCC CACCTAAGGC CATCTAGTGC CAACTCCCCC CCCACCATTC    2565

CCCTCACTGC ACTTTGGACC CCTGGGGCCA ACATCTCG                            2603
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 568 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Thr Ser Phe Ala Val Leu Met Asp Val Ser Arg Arg Glu Asn Gly
 1               5                  10                  15

Glu Ile Leu Pro Leu Lys Thr Leu Thr Tyr Val Ala Leu Gly Val Ala
            20                  25                  30

Leu Ala Ala Leu Leu Leu Thr Phe Phe Phe Leu Thr Leu Leu Arg Ile
        35                  40                  45

Leu Arg Ser Asn Gln His Gly Ile Arg Arg Asn Leu Thr Ala Ala Leu
    50                  55                  60

Gly Leu Ala Gln Leu Val Phe Leu Leu Gly Ile Asn Gln Ala Asp Leu
65                  70                  75                  80

Pro Phe Ala Cys Thr Val Ile Ala Ile Leu Leu His Phe Leu Tyr Leu
                85                  90                  95

Cys Thr Phe Ser Trp Ala Leu Leu Glu Ala Leu His Leu Tyr Arg Ala
            100                 105                 110

Leu Thr Glu Val Arg Asp Val Asn Thr Gly Pro Met Arg Phe Tyr Tyr
        115                 120                 125

Met Leu Gly Trp Gly Val Pro Ala Phe Ile Thr Gly Leu Ala Val Gly
    130                 135                 140

Leu Asp Pro Glu Gly Tyr Gly Asn Pro Asp Phe Cys Trp Leu Ser Ile
145                 150                 155                 160

Tyr Asp Thr Leu Ile Trp Ser Phe Ala Gly Pro Val Ala Phe Ala Val
                165                 170                 175

Ser Met Ser Val Phe Leu Tyr Ile Leu Ala Ala Arg Ala Ser Cys Ala
            180                 185                 190

Ala Gln Arg Gln Gly Phe Glu Lys Lys Gly Pro Val Ser Gly Leu Gln
        195                 200                 205

Pro Ser Phe Ala Val Leu Leu Leu Leu Ser Ala Thr Trp Leu Leu Ala
    210                 215                 220

Leu Leu Ser Val Asn Ser Asp Thr Leu Leu Phe His Tyr Leu Phe Ala
225                 230                 235                 240

Thr Cys Asn Cys Ile Gln Gly Pro Phe Ile Phe Leu Ser Tyr Val Val
                245                 250                 255

Leu Ser Lys Glu Val Arg Lys Ala Leu Lys Leu Ala Cys Ser Arg Lys
            260                 265                 270

Pro Ser Pro Asp Pro Ala Leu Thr Thr Lys Ser Thr Leu Thr Ser Ser
        275                 280                 285

Tyr Asn Cys Pro Ser Pro Tyr Ala Asp Gly Arg Leu Tyr Gln Pro Tyr
    290                 295                 300

Gly Asp Ser Ala Gly Ser Leu His Ser Thr Ser Arg Ser Gly Lys Ser
305                 310                 315                 320

Gln Pro Ser Tyr Ile Pro Phe Leu Leu Arg Glu Glu Ser Ala Leu Asn
                325                 330                 335
```

```
Pro Gly Gln Gly Pro Pro Gly Leu Gly Asp Pro Gly Ser Leu Phe Leu
            340                 345                 350

Glu Gly Gln Asp Gln Gln His Asp Pro Asp Thr Asp Ser Asp Ser Asp
        355                 360                 365

Leu Ser Leu Glu Asp Asp Gln Ser Gly Ser Tyr Ala Ser Thr His Ser
    370                 375                 380

Ser Asp Ser Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Ala Ala Phe
385                 390                 395                 400

Pro Gly Glu Gln Gly Trp Asp Ser Leu Leu Gly Pro Gly Ala Glu Arg
                405                 410                 415

Leu Pro Leu His Ser Thr Pro Lys Asp Gly Gly Pro Gly Pro Gly Lys
            420                 425                 430

Ala Pro Trp Pro Gly Asp Phe Gly Thr Thr Ala Lys Glu Ser Ser Gly
        435                 440                 445

Asn Gly Ala Pro Glu Glu Arg Leu Arg Glu Asn Gly Asp Ala Leu Ser
    450                 455                 460

Arg Glu Gly Ser Leu Gly Pro Leu Pro Gly Ser Ser Ala Gln Pro His
465                 470                 475                 480

Lys Gly Ile Leu Lys Lys Lys Cys Leu Pro Thr Ile Ser Glu Lys Ser
                485                 490                 495

Ser Leu Leu Arg Leu Pro Leu Glu Gln Cys Thr Gly Ser Ser Arg Gly
            500                 505                 510

Ser Ser Ala Ser Glu Gly Ser Arg Gly Gly Pro Pro Pro Arg Pro Pro
        515                 520                 525

Pro Arg Gln Ser Leu Gln Glu Gln Leu Asn Gly Val Met Pro Ile Ala
    530                 535                 540

Met Ser Ile Lys Ala Gly Thr Val Asp Glu Asp Ser Ser Gly Ser Glu
545                 550                 555                 560

Phe Leu Phe Phe Asn Phe Val His
                565
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCCAGGTCGA CCACATGACG AGCTTCGCT 29

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGGCTCTAGA GTTAATGCAC GAAGTTAAAG AAG 33

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTCATCTAGA AGTCCGCCAT CATGACGAGC AGCTTCGCTG TGCTC 45

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGGCTCTAGA GTTAATGCAC GAAGTTAAAG AAG 33

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTCATCTAGA AGTCCGCCAT CATGACGAGC TTCGCTGTGC TC 42

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGGCTCTAGA GTTAAGCGTA GTCTGGGACG TCGTATGGGT AATGCACGAA GGTTAAAGAA 60

G 61

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGGCTCTAGA GTTAATGCAC GAAGTTAAAG AAG 33

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GTCATCTAGA AGTCCGCCAT CATGACGAGC GGCGCTGTGC TC 42

(2) INFORMATION FOR SEQ .ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGGCTCTAGA GTTAAGCGTA GTCTGGGACG TCGTATGGGT AATGCACGAA GGTTAAAGAA 60

G 61

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGGCTCTAGA GTTAATGCAC GAAGTTAAAG AAG 33

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 439 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTGTGAAGTC GTCTTCCGCA ATGAGAGCCA CGTCAGCNGC CAGTGCAACC ACATGACGAG 60

CTTCGCTGTG CTCATGGACG TTTCTCGGCG GGAGANTGGG GAGATCCTGC CACTGAAGAC 120

ACTGACATAC GTGGCTNTAG GTGTCACCTT GGCTGCCCTT CTGCTCACCT TCTTCTTCCT 180

CACTCTCTTG CGTATCCTGC GCTCCAACCA ACACGGCATC CGACGTAACC TGACAGCTGC 240

CCTGGGCCTG GCTCAGCTGG TCTTCCTCCT GGGAATCAAC CAGGCTGACC TCCCTGTAAG 300

ATGCTCCTAC TGCCCAGAAA CTGTCCCCAC CTTCTCAGGC CGCCTCCCCA GCCCCCACTG 360

GCAACCCCTG CTCCTGCACC ATGAACTCTA ATAAGGTGCC TAGTGCAGAA CCTGGGCCAG 420

GGTTTCCTCT TCTGTGGCT 439

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 444 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CGNACAGTGA ACCTGTCCTT AGAAGACGAC CAGAGTGGCT CCTATGCCTC TACCCACTCA 60

TCAGACAGTG AGGAGGAAGA AGAGGAGGAG GAAGAGGAGG CCGCCTTCCC TGGAGAGCAG 120

GGCTGGGATA GCCTGCTGGG GCCTGGAGCA GAGAGACTGC CCCTGCACAG TACTCCCAAG 180

GATGGGGGCC CAGGGTCTGG NCAAGCCNCT NNGCCAGGNA GAACTTTGGG GACCACAGCA 240

AAAGAGAGTA GTGGCAACGN GGGCCCCTGG AGGAGCGGCT GCGNGGAGAA TGGAGATGCC 300

CTGTTTTCGA GAGGGGTTCC CTAGGGCCCC TTTCCAGGTT TTTTTGCCCA GCTTNACAAA 360

GGCATCTTTA AGAAGAATTT TTTNCCACCN TNANNGAGAG AGCAACTTCT NGGTTNCCCT 420

TGNGCAATTG AANAGGTTTT TCTC 444

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 517 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGCTGCTGGC ACTGCTCTCT GTCAACAGCG ACACCCTCCT CTTCCACTAC CTCTTTGCTA 60

CCTGCAATTG CATCCAGGGC CCCTTCATCT TCCTCTCCTA TGTGGTGCTT AGCAAGGAGG 120

TCCGGAAAGC ACTCAAGCTT GCCTGCAGCC GCAAGCCCAG CCCTGACCCT GCTCTGACCA 180

CCAAGTCCAC CCTGACCTCG TCCTACAACT GCCCCAGCCC CTACGCAGAT GGGCGGCTGT 240

ACCAGCCCTA ACGGAGACTC GGCCGGCTCT CTGCACAGCA CCAGTCGCTC GGGCAAGAGT 300

CAGCCCAGCT ACATCCCCTT CTTGCTGAGG GAGGAGTCCG CACTGAACCC TGGGCCAAGG 360

GCCNCCTGGG CTTGGGGGGA TTCCAGGAAG GCCTGTTTCC TGGGAAGGTT CAAAGACCAG 420

CAGCATTGAT TCCTGAACAA GGGATTTCGA CAGTGAACTT TTCTTTAGGA GGAGGACCAG 480

AGTNGGTTCT TATTGNTTTT AANCCATTTA TTNAGGA 517

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 385 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TTTTTTTTCT NCAAAAACTT TGTCTTGGAG ATGTTGGCCC CAGGGGTCCA AAGTGCAGTG 60

AGGGGAATGG TGGGGGGGGA GTTGGCACTA GATGGCCTTA GGTGGGGTGA GTACCCCTCC 120

CGGCAAGGCC CAGGCTCCTC AGATGGACGT TTCGTCGCTG CTGCGGCAGG CGATGGGGAG 180

CGGGGCAGGA TAAAGCACAA GACAGGAGCA GGGCATGAGT GCGGCTGGGG AAGGGAAGGG 240

AGCCTCGGGC TTAGGAACCA CGGCCCAGGG TTAATGCAGG AAGTTAAAGA AGAGAAATTC 300

GGAGCCTGAC GAGTCCTCAT CCACCGTGCC TGCCTTGATG CTCATGGCGA TGGGCATGAC 360

CCCGTTCAGC TGCTCCTGGG AGGCT 385

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 439 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CTGTGAAGTC GTCTTCCGCA ATGAGAGCCA CGTCAGCNGC CAGTGCAACC ACATGACGAG 60

CTTCGCTGTG CTCATGGACG TTTCTCGGCG GGAGANTGGG GAGATCCTGC CACTGAAGAC 120

ACTGACATAC GTGGCTNTAG GTGTCACCTT GGCTGCCCTT CTGCTCACCT TCTTCTTCCT 180

CACTCTCTTG CGTATCCTGC GCTCCAACCA ACACGGCATC CGACGTAACC TGACAGCTGC 240

CCTGGGCCTG GCTCAGCTGG TCTTCCTCCT GGGAATCAAC CAGGCTGACC TCCCTGTAAG 300

ATGCTCCTAC TGCCCAGAAA CTGTCCCCAC CTTCTCAGGC CGCCTCCCCA GCCCCCACTG 360

GCAACCCCTG CTCCTGCACC ATGAACTCTA ATAAGGTGCC TAGTGCAGAA CCTGGGCCAG 420

GGTTTCCTCT TCTGTGGCT 439

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 321 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGANTCGTCN ATCGTGGTTG CACTGGCAGC TGACGTGGCT CTCATTGCGG AAGACGACTT 60

CACAGCCTCT GNCCGACCAG CCACCTGTGC CACTGACCAG GATTGAATGG TTCCAGAAGA 120

CACAGATGGG CTTGGTCCGC TCCTCTGTCT CCAGCAGGCG GAACTGCACC GTGACGGGTT 180

TGTCCAGGGC CCGGGGCAGA AGCTCCTCAT CATCATGGAC GCTGATGCTC ACCACGGGTG 240

TGTTGATGAT CGGGCGTTTG GGGACTCTCA AGCTGCGCTT GTCANGGGTC ATAGTTATGA 300

GGCANGTAGC CCGGCCAGGG T 321

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 300 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCGACCGCAC CCTGGCCGGG CTACTGCTAC ATAACTATGA CCCTGACAAG CGCATTGACN 60

ACGTCCCCAA ACGCCCGATC ATCAACACAC CCGTGGTGAG CATCAGCGTC CATGATGATG 120

AGGAGCTTCT GCCCCGGGCC CTGGACAAAC CCGTCACGGT GCAGTTCCGC CTGCTGGAGA 180

CAGAGGAGCG GACCAAGCCC ATCTGTGTCT TCTGGGAACC ATTCAATCCT GGTCAGTGGC 240

ACAGGTGGCT GGTCGGCCAG AGNTCGTCGA AGTCGTCTTC CGCAATGAGA GCCACGTCAG 300

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 239 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TTTCTGAAAA ACTTTNTCTT GGAGATGTTG GCCCCAGGGG TCCAAAGTGC AGTNAGGGGA 60

ATGGTGGGGG GGGAGTTGGC ACTAGATGGC CTTAGGTGGG GTGAGTACCC CTCCCGGNAA 120

GGCCCAGGCT CCTNAAATGG ACGTTTCGTC GCTGTCGGAG CCTGACGAGT CCTCATCCAC 180

CGNGACTGCC TTGATGCTCA TNGCGATGGG CATNACCCCG TTNAGCTGCT CCTGAGAGG 239

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 170 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCAGGGTGTG TTTTCCCACC AGCCGCAGAG AGCCAGGATG GACGTTCCTC GGACGGACGG 60

TTTTCCTGCT TGGGAATGTT CCTGGGCTGT GAGATCCACT CTTCTGGGCA GGTGGTTAGC 120

ACCTAACGTT TTTCCCTCAC TTCCCCCCAA ATTCTTAAGT CCTTTGGTCC 170

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 280 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AAAGGACTTA AGAATTTGGG GGGAAGTGAG GGAAAAACGT TAGGTGCTAA CCACCTGCCC 60

AGAAGAGTGG ATCTCACAGC CCAGGAACAT TCCCAAGCAG GAAAACCGTC CGTCCGAGGA 120

ACGTCCATCC TGGCTCTCTG CGGCTGGTGG GAAAACACAC CCTGCCCTGN GGGGCTGTCC 180

AGGNCTTCTC CCCCACACCC TCAGGCCGAG ATCCGTGAGA GACCCACTTT GCTCCAACAA 240

CTTGAAACAA GTCACTTTAC CCTCCTTAGG ACCCATTTTG 280

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 138 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TGGACCAAAG GACTTAAGAA TTTGCCGGGA AGTGAGGGAA AAACGTTAGT GCTAACACCG 60

CTGCCCAGAA GAGTGGATCT CACAGCCCAG GAACATTCCC AAGCAGGAAA ACCGTCCGTC 120

CGAGGAACGT CCATCCTG 138

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 236 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TNAGAATTTG GGGGGAAGTG AGGGAAAACG TTAGGTGCTA ACACCGTGCC CAGAAGAGTG 60

GATCTCACAG CCCAGGAACA TTCCAAGCAG GANAACCGTC CGTCCGAGGA ACGTCCATCC 120

TGGCTCTCTG CGGCTGGTGG GAAAACACAN CCTGCCCTGA GGCTGTCCAG GCCTTCTCCC 180

CCACACCCTC AGGCCGAGAT CCGTGAGAGA CCCANTTTGN TCCAACAACT TGAAAC 236

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 201 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AACGGGGAGG TCCTGCCTCT GAAGATTGTC ACCTATGCCG CTGTGTCCTT GTCACTGGCA 60

GCCCTGCTGG TGGCCTTCGT CCTCCTGAGC CTGGTCCGCA TGCTGCGCTC CAACCTGCAC 120

AGCATTCACA AGCACCTCGC CGTGGCGCTC TTCCTCTCTC AGCTGGTGTT CGTGATTGGG 180

ATCAACCAGA CGGAAAACCC G 201

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 292 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ACCCCACCTA AGGCCATCTA GTGCCAACTC CCCCCCCACC ATTCCCCTCA CTGCACTTTG 60

GACCCCTGGG GCCAACATCT CCAAGACAAA GTTTTTCAGA AAAGAGGAAA AAAAGAATTT 120

AAAAAAGGAT CTCCACTCTT CATGACTTCA GGGATTCATT TTTTTTATAC GCTGGAAATT 180

GACTCCCCTT TCCCTTCCCA AAGAGGATAG GACCTCCCAG GATGCTTCCC AGCCTCTCCT 240

CAGTTTCCCA TCTGCTGTGC CTCTGGGAGG AGAGGGACTC CTGGGGGGCC TG 292

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 4, line 13.

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on an additional sheet ☐

Name of depositary institution

American Type Culture Collection

Address of depositary institution *(including postal code and country)*

12301 Parklawn Drive
Rockville, Maryland 20852
United States of America

Date of deposit	Accession Number
September 23, 1996	ATCC 97730

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* — This information is continued on an additional sheet ☐

DNA plasmid, HCEPR64

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

Form PCT/RO/134 (July 1992)

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a nucleotide sequence encoding the calcitonin receptor polypeptide having the amino acid sequence from about 1 to about 568 in SEQ ID NO:2;

(b) a nucleotide sequence encoding the calcitonin receptor polypeptide having the amino acid sequence from about 2 to about 568 in SEQ ID NO:2;

(c) a nucleotide sequence encoding the calcitonin receptor polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97730;

(d) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor extracellular regions selected from the group consisting of: amino acid residues from about 1 to about 18, amino acid residues from about 148 to about 159, amino acid residues from about 188 to about 209, and amino acid residues from about 231 to about 236 in SEQ ID NO:2;

(e) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor intracellular regions selected from the group consisting of:amino acid residues from about 50 to about 59, amino acid residues from about 110 to about 125, amino acid residues from about 188 to about 209, and amino acid residues from about 259 to about 560 in SEQ ID NO:2;

(f) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 1 (amino acid residues from about 20 to about 49 in SEQ ID NO:2);

(g) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 2 (amino acid residues from about 60 to about 80 in SEQ ID NO:2);

(h) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 3 (amino acid residues from about 82 to about 109 in SEQ ID NO:2);

(i) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 4 (amino acid residues from about 126 to about 147 in SEQ ID NO:2);

(j) a nucleotide sequence encoding the calcitonin receptor transmembrane domain 5 (amino acid residues from about 160 to about 187 in SEQ ID NO:2);

(k) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 6 (amino acid residues from about 209 to about 230 in SEQ ID NO:2);

(l) a nucleotide sequence encoding a polypeptide comprising the calcitonin receptor transmembrane domain 7 (amino acid residues from about 237 to about 258 in SEQ ID NO:2); and

(m) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), (h), (i), (j), (k) or (l).

2. An isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of an epitope-bearing portion of an calcitonin receptor polypeptide having an amino acid sequence in (a), (b), (c), (d), (e), (f), (g), (h), (i), (j), (k) or (l) of claim 1.

3. The isolated nucleic acid molecule of claim 2, which encodes an epitope-bearing portion of any of the calcitonin receptor polypeptide selected from the group consisting of: a polypeptide comprising amino acid residues from about 10 to about 20 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 49 to about 60 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 113 to about 123 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 145 to about 154 in SEQ ID NO:2; a polypeptide comprising

amino acid residues from about 189 to about 209 in SEQ ID NO:2; and a polypeptide comprising amino acid residues from about 259 to about 560 in SEQ ID NO:2.

4. An isolated nucleic acid molecule, comprising a polynucleotide having a sequence slected from the group consisting of:

(a) the nucleotide sequence of a fragment of the sequence shown in SEQ ID NO:1, wherein said fragment comprises at least 50 contiguous nucleotides of SEQ ID NO:1, provided that said isolated nucleic acid molecule is not HCEPR64R (SEQ ID NO:13), HCETE57R (SEQ ID NO:14), R60111 (SEQ ID NO:15), R60049 (SEQ ID NO:16), AA326354 (SEQ ID NO:17), W52864 (SEQ ID NO:18), AA037392 (SEQ ID NO:19), T06233 (SEQ ID NO:20), AA477598 (SEQ ID NO:21), F13341 (SEQ ID NO:22), AA477470 (SEQ ID NO:23), T77092 (SEQ ID NO:24), H55526 (SEQ ID NO:25), AA349458 (SEQ ID NO:26). or any subfragment thereof; and

(b) a nucleotide sequence complementary to a nucleotide sequence in (a).

5. A method for making a recombinant vector comprising inserting an isolated nucleic acid molecule of claim 1 into a vector.

6. A recombinant vector produced by the method of claim 5.

7. A method of making a recombinant host cell comprising introducing the recombinant vector of claim 6 into a host cell.

8. A recombinant host cell produced by the method of claim 7.

9. A recombinant method for producing any of the calcitonin receptor polypeptide, comprising culturing the recombinant host cell of claim 8 under conditions such that said polypeptide is expressed and recovering said polypeptide.

10. An isolated calcitonin receptor polypeptide having an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) the amino acid sequence of the calcitonin receptor polypeptide comprising the amino acid sequence from about 1 to about 568 in SEQ ID NO:2;

(b) the amino acid sequence of the calcitonin receptor polypeptide comprising the amino acid sequence from about 2 to about 568 in SEQ ID NO:2;

(c) the amino acid sequence of the calcitonin receptor polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97730;

(d) the amino acid sequence of the calcitonin receptor extracellular regions selected from the group consisting of: amino acid residues from about 1 to about 18, amino acid residues from about 148 to about 159, amino acid residues from about 188 to about 209, and amino acid residues from about 231 to about 236 in SEQ ID NO:2;

(e) the amino acid sequence of the calcitonin receptor intracellular regions selected from the group consisting of:amino acid residues from about 50 to about 59, amino acid residues from about 110 to about 125, amino acid residues from about 188 to about 209, and amino acid residues from about 259 to about 560 in SEQ ID NO:2;

(f) the amino acid sequence of the calcitonin receptor transmembrane domain 1 (amino acid residues from about 20 to about 49 in SEQ ID NO:2);

(g) the amino acid sequence of the calcitonin receptor transmembrane domain 2 (amino acid residues from about 60 to about 80 in SEQ ID NO:2);

(h) the amino acid sequence of the calcitonin receptor transmembrane domain 3 (amino acid residues from about 82 to about 109 in SEQ ID NO:2);

(i) the amino acid sequence of the calcitonin receptor transmembrane domain 4 (amino acid residues from about 126 to about147 in SEQ ID NO:2);

(j) the amino acid sequence of the calcitonin receptor intracellular domain 5 (amino acid residues from about 160 to about 187 in SEQ ID NO:2);

(k) the amino acid sequence of the calcitonin receptor transmembrane domain 6 (amino acid residues from about 209 to about 230 in SEQ ID NO:2);

(l) the amino acid sequence of the calcitonin receptor transmembrane domain 7 (amino acid residues from about 237 to about 258 in SEQ ID NO:2); and

(m) the amino acid sequence of an epitope-bearing portion of any one of the peptides of (a), (b), (c), (d), (e), (f), (g), (h), (i), (j), (k) or (l).

11. An isolated polypeptide comprising an epitope-bearing portion of any of the calcitonin receptor, wherein said portion is selected from the group consisting of: a polypeptide comprising amino acid residues from about 10 to about 20 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 49 to about 60 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 113 to about 123 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 145 to about 154 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about 189 to about 209 in SEQ ID NO:2; and a polypeptide comprising amino acid residues from about 259 to about 560 in SEQ ID NO:2.

12. An isolated antibody that binds specifically to a calcitonin receptor polypeptide of claim 10.

13. An isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide having a nucleotide sequence identical to a nucleotide sequence in (a), (b), (c), (d) (e), (f), (g), (h), (i), (j), (k), (l) or (m) in claim 1 wherein said polynucleotide which hybridizes does not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues.

14. An isolated nucleic acid molecule comprising a polynucleotide encoding an calcitonin receptor polypeptide wherein, except for at least one to fifty conservative amino acid substitutions, said polypeptide has a sequence selected from the group consisting of:

(a) the amino acid sequence of the calcitonin receptor polypeptide comprising the amino acid sequence from about 1 to about 568 in SEQ ID NO:2;

(b) the amino acid sequence of the calcitonin receptor polypeptide comprising the amino acid sequence from about 2 to about 568 in SEQ ID NO:2;

(c) the amino acid sequence of the calcitonin receptor polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 97730;

(d) the amino acid sequence of the calcitonin receptor extracellular regions selected from the group consisting of: amino acid residues from about 1 to about 18, amino acid residues from about 148 to about 159, amino acid residues from about 188 to about 209, and amino acid residues from about 231 to about 236 in SEQ ID NO:2;

(e) the amino acid sequence of the calcitonin receptor intracellular regions selected from the group consisting of:amino acid residues from about 50 to about 59, amino acid residues from about 110 to about 125, amino acid residues from about 188 to about 209, and amino acid residues from about 259 to about 560 in SEQ ID NO:2;

(f) the amino acid sequence of the calcitonin receptor transmembrane domain 1 (amino acid residues from about 20 to about 49 in SEQ ID NO:2);

(g) the amino acid sequence of the calcitonin receptor transmembrane domain 2 (amino acid residues from about 60 to about 80 in SEQ ID NO:2);

(h) the amino acid sequence of the calcitonin receptor transmembrane domain 3 (amino acid residues from about 82 to about 109 in SEQ ID NO:2);

(i) the amino acid sequence of the calcitonin receptor transmembrane domain 4 (amino acid residues from about 126 to about147 in SEQ ID NO:2);

(j) the amino acid sequence of the calcitonin receptor intracellular domain 5 (amino acid residues from about 160 to about 187 in SEQ ID NO:2);

(k) the amino acid sequence of the calcitonin receptor transmembrane domain 6 (amino acid residues from about 209 to about 230 in SEQ ID NO:2);

(l) the amino acid sequence of the calcitonin receptor transmembrane domain 7 (amino acid residues from about 237 to about 258 in SEQ ID NO:2); and

(m) the amino acid sequence of an epitope-bearing portion of any one of the peptides of (a), (b), (c), (d), (e), (f), (g), (h), (i), (j), (k) or (l).

15. An isolated antibody that binds specifically to a calcitonin receptor polypeptide of claim 14.

16. A method of treating abnormal conditions related to an excess of calcitonin receptor activity comprising administering to a patient in need there of an antagonist of the polypeptide of claim 14.

17. A method of treating abnormal conditions related to an under expression of calcitonin receptor activity comprising administering to a patient in need there of an angonist of the polypeptide of claim 14.

18. A method of treating diseases or disorders selected from the group consisting of bacterial, fungal, protozoan and viral infections; infection caused by HIV-1 or HIV-2; pain; cancers; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation; and dyskinesias comprising administering to a patient in need thereof an effective amount of an antagonist of the the polypeptide of claim 14.

```
   1 CCTTTCCCACCAGCCGCAGAGAGCCAGGATGGACGTTCCTCGGACGGACGGTTTTCCTGC  60
  61 TTGGGAATGTTCCTGGGCTGTGAGATCCACTCTTCTGGGCAGGTGGTTAGCACCTAACGT 120
 121 TTTTCCCTCACTTCCCCCCAAATTCTTAAGTCCTTTGGTCCATTTCACTGCTCGGACCTT 180
 181 GAGACAACAGTCATTCTGCCTGAGTCTGTCTTCAGAGAGACGCCCCCCGTGGTCAGGCCC 240
 241 GCAGGCCCCGGAGAGGCCCAGGAGCCAGAGGAGCTGGCACGGCGACAGCGACGGCACCCG 300
 301 GAGCTGAGCCAGGGTGAGGCTGTGGCCAGCGTCATCATCTACCGCACCCTGGCCGGGCTA 360
 361 CTGCCTCATAACTATGACCCTGACAAGCGCAGCTTGAGAGTCCCCAAACGCCCGATCATC 420
 421 AACACACCCGTGGTGAGCATCAGCGTCCATGATGATGAGGAGCTTCTGCCCCGGGCCCTG 480
 481 GACAAACCCGTCACGGTGCAGTTCCGCCTGCTGGAGACAGAGGAGCGGACCAAGCCCATC 540
 541 TGTGTCTTCTGGAACCATTCAATCCTGGTCAGTGGCACAGGTGGCTGGTCGGCCAGAGGC 600
 601 TGTGAAGTCGTCTTCCGCAATGAGAGCCACGTCAGCTGCCAGTTCAACCACATGACGAGC 660
 -16                                                       M T S   3
 661 TTCGCTGTGCTCATGGACGTTTCTCGGCGGGAGAATGGGGAGATCCTGCCACTGAAGACA 720
   4 F A V L M D V S R R E N G E I L P L K T                      23
 721 CTGACATACGTGGCTCTAGGTGTCGCCTTGGCTGCCCTTCTGCTCACCTTCTTCTTCCTC 780
  24 L T Y V A L G V A L A A L L L T F F F L                      43
 781 ACTCTCTTGCGTATCCTGCGCTCCAACCAACACGGCATCCGACGTAACCTGACAGCTGCC 840
  44 T L L R I L R S N Q H G I R R N L T A A                      63
 841 CTGGGCCTGGCTCAGCTGGTCTTCCTCCTGGGAATCAACCAGGCTGACCTCCCTTTTGCC 900
  64 L G L A Q L V F L L G I N Q A D L P F A                      83
 901 TGCACAGTCATTGCCATCCTGCTGCACTTCCTGTACCTCTGCACCTTTTCCTGGGCTCTG 960
  84 C T V I A I L L H F L Y L C T F S W A L                     103
 961 CTGGAGGCCTTGCACCTGTACCGGGCACTCACTGAGGTGCGCGATGTCAACACCGGCCCC 1020
 104 L E A L H L Y R A L T E V R D V N T G P                     123
1021 ATGCGCTTCTACTACATGCTGGGCTGGGGCGTGCCTGCCTTCATCACAGGGCTAGCCGTG 1080
 124 M R F Y Y M L G W G V P A F I T G L A V                     143
1081 GGCCTGGACCCCGAGGGCTACGGGAACCCTGACTTCTGCTGGCTCTCCATCTATGACACG 1140
 144 G L D P E G Y G N P D F C W L S I Y D T                     163
1141 CTCATCTGGAGTTTTGCTGGCCCGGTGGCCTTTGCCGTCTCGATGAGTGTCTTCCTGTAC 1200
 164 L I W S F A G P V A F A V S M S V F L Y                     183
1201 ATCCTGGCGGCCCGGGCCTCCTGTGCTGCCCAGCGGCAGGGCTTTGAGAAGAAAGGTCCT 1260
 184 I L A A R A S C A A Q R Q G F E K K G P                     203
1261 GTCTCGGGCCTGCAGCCCTCCTTCGCCGTCCTCCTGCTGCTGAGCGCCACGTGGCTGCTG 1320
 204 V S G L Q P S F A V L L L L S A T W L L                     223
1321 GCACTGCTCTCTGTCAACAGCGACACCCTCCTCTTCCACTACCTCTTTGCTACCTGCAAT 1380
 224 A L L S V N S D T L L F H Y L F A T C N                     243
1381 TGCATCCAGGGCCCCTTCATCTTCCTCTCCTATGTGGTGCTTAGCAAGGAGGTCCGGAAA 1440
 244 C I Q G P F I F L S Y V V L S K E V R K                     263
1441 GCACTCAAGCTTGCCTGCAGCCGCAAGCCCAGCCCTGACCCTGCTCTGACCACCAAGTCC 1500
 264 A L K L A C S R K P S P D P A L T T K S                     283
1501 ACCCTGACCTCGTCCTACAACTGCCCCAGCCCCTACGCAGATGGGCGGCTGTACCAGCCC 1560
 284 T L T S S Y N C P S P Y A D G R L Y Q P                     303
1561 TACGGAGACTCGGCCGGCTCTCTGCACAGCACCAGTCGCTCGGGCAAGAGTCAGCCCAGC 1620
 304 Y G D S A G S L H S T S R S G K S Q P S                     323
1621 TACATCCCCTTCTTGCTGAGGGAGGAGTCCGCACTGAACCCTGGCCAAGGGCCCCCTGGC 1680
 324 Y I P F L L R E E S A L N P G Q G P P G                     343
1681 CTGGGGGATCCAGGCAGCCTGTTCCTGGAAGGTCAAGACCAGCAGCATGATCCTGACACG 1740
 344 L G D P G S L F L E G Q D Q Q H D P D T                     363
1741 GACTCCGACAGTGACCTGTCCTTAGAAGACGACCAGAGTGGCTCCTATGCCTCTACCCAC 1800
 364 D S D S D L S L E D D Q S G S Y A S T H                     383
1801 TCATCAGACAGTGAGGAGGAAGAAGAGGAGGAGGAAGAGGAGGCCGCCTTCCCTGGAGAG 1860
 384 S S D S E E E E E E E E E E A A F P G E                     403
1861 CAGGGCTGGGATAGCCTGCTGGGGCCTGGAGCAGAGAGACTGCCCCTGCACAGTACTCCC 1920
 404 Q G W D S L L G P G A E R L P L H S T P                     423
1921 AAGGATGGGGGCCCAGGGCCTGGCAAGGCCCCCTGGCCAGGAGACTTTGGGACCACAGCA 1980
 424 K D G G P G P G K A P W P G D F G T T A                     443
1981 AAAGAGAGTAGTGGCAACGGGGCCCCTGAGGAGCGGCTGCGGGAGAATGGAGATGCCCTG 2040
 444 K E S S G N G A P E E R L R E N G D A L                     463
2041 TCTCGAGAGGGGTCCCTAGGCCCCCTTCCAGGCTCTTCTGCCCAGCCTCACAAAGGCATC 2160
 464 S R E G S L G P L P G S S A Q P H K G I                     483
2101 CTTAAGAAGAAGTGTCTGCCCACCATCAGCGAGAAGAGCAGCCTCCTGCGGCTCCCCCTG 2160
 484 L K K K C L P T I S E K S S L L R L P L                     503
```

FIG.1A

```
2161  GAGCAATGCACAGGGTCTTCCCGGGGCTCCTCCGCTAGTGAGGGCAGCCGGGGCGGCCCC  2220
 504  E  Q  C  T  G  S  S  R  G  S  S  A  S  E  G  S  R  G  G  P    523
2221  CCTCCCCGCCCACCGCCCCGGCAGAGCCTCCAGGAGCAGCTGAACGGGGTCATGCCCATC  2280
 524  P  P  R  P  P  P  R  Q  S  L  Q  E  Q  L  N  G  V  M  P  I    543
2281  GCCATGAGCATCAAGGCAGGCACGGTGGATGAGGACTCGTCAGGCTCCGAATTTCTCTTC  2340
 544  A  M  S  I  K  A  G  T  V  D  E  D  S  S  G  S  E  F  L  F    563
2341  TTTAACTTCGTGCATTAACCCTGGGCCGTGGTTCCTACGCCCGAGGCTCCCTTCCCTTCC  2400
 564  F  N  F  V  H  *                                              583
2401  CCAGCCGCACTCATGCCCTGCTCCTGTCTTGTGCTTTATCCTGCCCCGCTCCCCATCGCC  2460
2461  TGCCCGCAGCAGCGACGAAACGTCCATCTGAGGAGCCTGGGCCTTGCCGGGAGGGGTACT  2520
2521  CACCCCACCTAAGGCCATCTAGTGCCAACTCCCCCCCCACCATTCCCCTCACTGCACTTT  2580
2581  GGACCCCTGGGGCCAACATCTCG  2603
```

FIG.1B

```
1    MTSFAVLMDVSRRENGEILPLKTLTYVALGVALAALLLTFFFLTLLRILR
                              TM1
51   SNQHGIRRNLTAALGLAQLVFLLGINQADLPFACTVIAILLHFLYLCTFS
                    TM2                TM3
101  WALLEALHLYRALTEVRDVNTGPMRFYYMLGWGVPAFITGLAVGLDPEGY
                                    TM4
151  GNPDFCWLSIYDTLIWSFAGPVAFAVSMSVFLYILAARASCAAQRQGFEK
                          TM5
201  KGPVSGLQPSFAVLLLLSATWLLALLSVNSDTLLFHYLFATCNCIQGPFI
                    TM6                      TM7
251  FLSYVVLSKEVRKALKLACSRKPSPDPALTTKSTLTSSYNCPSPYADGRL

301  YQPYGDSAGSLHSTSRSGKSQPSYIPFLLREESALNPGQGPPGLGDPGSL

351  FLEGQDQQHDPDTDSDSDLSLEDDQSGSYASTHSSDSEEEEEEEEEEAAF

401  PGEQGWDSLLGPGAERLPLHSTPKDGGPGPGKAPWPGDFGTTAKESSGNG

451  APEERLRENGDALSREGSLGPLPGSSAQPHKGILKKKCLPTISEKSSLLR

501  LPLEQCTGSSRGSSASEGSRGGPPPRPPPRQSLQEQLNGVMPIAMSIKAG

551  TVDEDSSGSEFLFFNFVH*
```

FIG.2

PCT/US 97/21330

Applicant's or agent's file reference number	International application No.
1488.066PC01	TBA

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 4, line 13

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on an additional sheet ☐

Name of depositary institution

American Type Culture Collection

Address of depositary institution *(including postal code and country)*

12301 Parklawn Drive
Rockville, Maryland 20852
United States of America

Date of deposit	Accession Number
September 23, 1996	ATCC 97730

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* — This information is continued on an additional sheet ☐

DNA plasmid, HCEPR64
In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer: Yvette E. Simms	Authorized officer

Form PCT/RO/134 (July 1992)

(DNA plasmid, HCEPR64)

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

DENMARK

The applicant hereby request that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

SWEDEN

The applicant hereby request that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby request that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

NETHERLANDS

The applicant hereby request that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INTERNATIONAL SEARCH REPORT

Intern nal Application No
PCT/US 97/21330

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C07K14/72 //C07K14/585

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NAGASE T ET AL.: "Prediction of the coding sequences of unidentified human genes. VI. " DNA RESEARCH, vol. 3, no. 5, 31 October 1996, TOKYO JP, pages 321-329, XP002059454 see the whole document & EMBL SEQUENCE DATABASE, HEIDELBERG, DE, ACCESSION NR.: D87469, 9 NOVEMBER 1996, NAGASE,T. "Human mRNA for KIAA0279 gene" see abstract	4
Y		1-3

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search
19 March 1998

Date of mailing of the international search report
07.04.98

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer
De Kok, A

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 97/21330

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ADAMS M D ET AL: "INITIAL ASSESSMENT OF HUMAN GENE DIVERSITY AND EXPRESSION PATTERNS BASED UPON 83 MILLION NUCLEOTIDES OF CDNA SEQUENCE" NATURE, vol. 377, no. supp, 28 September 1995, LONDON GB, pages 3-174, XP002042918 see page 3 - page 17 ---	1-3
X	EMBL SEQUENCE DATABASE, HEIDELBERG, DE, ACCESSION NR.: F13341, 5 March 1995, Genethon, "Human cDNA clone c-3ma05" XP002059456 see abstract ---	4
A	MUFF R ET AL.: "Calcitonin, calcitonin gene-related peptide, adrenomedullin and amylin: Homologous peptides, separate receptors and overlapping biological actions" EUROPEAN JOURNAL OF ENDOCRINOLOGY, vol. 133, no. 1, 1995, OSLO NO, pages 17-20, XP002059455 see the whole document ---	1-18
A	WO 96 05221 A (HUMAN GENOME SCIENCES INC) 22 February 1996 cited in the application see the whole document ---	1-18
A	WO 94 08006 A (ZYMOGENETICS INC) 14 April 1994 see the whole document ---	1-18
A	WO 93 10149 A (THE GENERAL HOSPITAL CORP) 27 May 1993 see the whole document -----	1-18

1

INTERNATIONAL SEARCH REPORT

In ...national application No.
PCT/US 97/21330

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

see FURTHER INFORMATION sheet PCT/ISA/210

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Remark : Although claims 16-18 , are directed to a method of treatment of the human/animal body , the search has been carried out and based on the alleged effects of the compound/composition, insofar as these compounds/compositions has been technically disclosed (Art.6 PCT)

INTERNATIONAL SEARCH REPORT

information on patent family members

Inter nal Application No
PCT/US 97/21330

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9605221 A	22-02-96	AU 1800295 A	07-03-96
		AU 7953194 A	07-03-96
		EP 0777684 A	11-06-97
		EP 0777493 A	11-06-97
		WO 9604928 A	22-02-96
WO 9408006 A	14-04-94	CA 2145866 A	14-04-94
		EP 0663006 A	19-07-95
		JP 8501942 T	05-03-96
		US 5683884 A	04-11-97
		US 5674981 A	07-10-97
		US 5622839 A	22-04-97
		US 5674689 A	07-10-97
WO 9310149 A	27-05-93	US 5516651 A	14-05-96
		CA 2123582 A	27-05-93
		EP 0674656 A	04-10-95